

Réalisé par Agraffel Nicolas

Tuteur Universitaire : Jean-Christophe Chasle

Mémoire Master 1 MBFA – Parcours ingénierie économique et financière –
Année 2023-2024

*La prise en compte des enjeux extra financiers dans le
monde de la finance.*

Problématique : Les facteurs ESG sont-ils pertinents dans la prédiction de la performance financière en analyse quantitative ? Etude empirique sur des entreprises du CAC40.

Sommaire :

- Introduction
- 1) Les facteurs ESG : Définition, cadre légal et leurs impacts théorique sur la performance des entreprises.
 - a) Définition et sélection des critères ESG pour l'analyse quantitative :
 - b) Cadre légal, normes et labels ESG : Synthèse et lien avec la performance financière.
- 2) L'analyse quantitative : Implémentation des variables et création du modèle.
 - a) Intégration des variables dans le modèle.
 - b) Création du modèle et analyse des résultats.
- Conclusion
- Limites et ouvertures
- Table des matières
- Bibliographie

Introduction :

«Sustainability is not a charity; it's good business» Paul Polman, ancien PDG d'Unilever

Depuis plusieurs décennies, le monde des affaires a été témoin d'une évolution significative des critères de performance des entreprises. Traditionnellement, les mesures de performance se concentraient principalement sur des indicateurs financiers tels que le chiffre d'affaires, le bénéfice, ou le retour sur investissement. Cependant, une prise de conscience croissante autour des enjeux environnementaux, sociaux et de gouvernance, ESG (Environnement, Social, Gouvernance) a émergé. Elle est en grande partie due à des préoccupations mondiales croissantes concernant le changement climatique, la justice sociale, ou encore la transparence et la responsabilité des entreprises.

L'intégration des critères ESG dans l'analyse de la performance des entreprises représente une évolution majeure. Cela pose la question de savoir pourquoi ces facteurs non financiers impactent la performance financière des entreprises. La pertinence de l'ajout des critères ESG dans les modèles d'analyse quantitative est aujourd'hui largement débattue, avec des arguments allant de l'amélioration de la gestion des entreprises face aux risques extrêmes, à la création de valeur à long terme pour les actionnaires et les parties prenantes.

Ce mémoire propose d'explorer cette problématique en se concentrant spécifiquement sur les entreprises du CAC40, indice de référence des principales entreprises cotées en France. Les entreprises de cet indice, par leur taille et leur visibilité, offrent un terrain d'étude pertinent. Nous chercherons à déterminer si les critères ESG sont pertinents à utiliser dans des modèles quantitatifs, dans l'objectif de prédire la performance financière future.

Pour ce faire, cette étude a été structurée en deux grandes parties. La première est une partie plutôt théorique, dédiée à une revue approfondie des facteurs ESG : nous commencerons par définir et sélectionner les critères ESG les plus pertinents pour notre analyse quantitative, en nous appuyant sur la littérature existante et les standards internationaux. Nous examinerons ensuite le cadre légal et les normes actuelles entourant les critères ESG, ainsi que les labels et certifications qui influencent la transparence et la conformité des entreprises aux exigences ESG. Cette section vise à établir un lien clair entre ces critères ESG et leur impact sur la performance financière des entreprises.

La deuxième partie de ce travail se concentre sur l'aspect purement quantitatif de l'étude. Nous détaillerons comment les variables ESG sont intégrées dans le modèle d'analyse quantitatif, en expliquant les choix méthodologiques et les techniques statistiques utilisées pour utiliser les données. Nous présenterons ensuite la création du modèle en lui-même, en discutant des méthodes d'estimation employées. Puis, nous analyserons les résultats obtenus, en mettant en lumière les éventuels impacts positifs ou négatifs des facteurs ESG sur la prédiction de la performance financière des entreprises du CAC40.

Néanmoins, pour réaliser cette étude, certaines hypothèses doivent être posées au préalable. Tout d'abord, il faut supposer que les facteurs ESG permettent de façon sûre, d'améliorer la performance des entreprises. Beaucoup d'étude vont en effet dans ce sens, et l'explication du lien entre les critères et la performance sera détaillé plus tard. Cependant, le manque de recul face à l'ESG, du à l'apparition plutôt récente des critères ne permettent pas d'assurer à 100% cette hypothèse. Également, il faut partir du principe que les données ESG sont correctement mesurées et recueillies. La source de ces données est fiable, et les entreprises sont soumises à des réglementations strictes en matière de reporting, point qui sera également détaillé plus tard. Mais, les mesures des facteurs ESG ont déjà été modifié par le passé, et si un nouveau changement été opéré, l'analyse porté dans ce travail pourrait de fait, être faussée. Ainsi, seul l'approbation de ces deux hypothèses permettent de réaliser ce travail dans les meilleures conditions.

Cette étude cherche non seulement à contribuer à la compréhension des relations entre les facteurs ESG et la performance financière, mais aussi de montrer comment utiliser ces variables dans des modèles quantitatif et s'il est utile de le faire. Elle n'a cependant pas pour but de créer un modèle véritablement efficace et précis dans la prédiction. Des choix d'analyse seront ainsi faits pour permettre une meilleure étude du sujet, mais ils entraineront un potentiel pouvoir prédictif inférieur.

1 – Les facteurs ESG : Définition, cadre légal et leurs impacts théorique sur la performance des entreprises.

A) Définition et sélection des critères ESG pour l'analyse quantitative :

« Les critères extra-financiers permettent d'évaluer un acteur économique, par exemple une société cotée, en dehors des critères financiers habituels (chiffre d'affaires, prix de l'action, perspectives de croissance, etc.). Il s'agit de prendre en compte son impact et ses performances via un autre prisme, comme celui de l'environnement.

En matière de finance durable, ces critères sont appelés ESG pour Environnemental, Social et de Gouvernance. Ils font l'objet d'une notation par des agences spécialisées ou non : plus une société est performante sur un ou plusieurs critères, meilleure est sa note. Et c'est sur la base de ces notations notamment que les sociétés de gestion peuvent choisir les actifs qui composent leurs fonds d'investissement responsable par exemple. » [Qu'est-ce qu'un critère extra-financier ? | AMF \(amf-france.org\)](#)

Le **critère environnemental** prend en compte l'impact de la société sur l'environnement, grâce à des données tel que l'émission de gaz à effet de serre, la consommation d'énergie etc...

Ce critère concerne toutes les entreprises, que leurs productions soient orientées vers l'écologie ou non. L'entreprise sera alors notée en fonction de l'efficacité et de l'importance des actions qu'elle met en place pour être plus respectueuse de l'environnement. Elle peut par exemple modifier sa manière de produire pour baisser l'émission de gaz à effet de serre.

Le **critère social** est quant à lui plus large. Il peut par exemple prendre en compte les éléments internes à la société, comme l'accessibilité pour les travailleurs handicapés ou l'égalité salariale. Mais il concerne aussi les parties prenantes de l'entreprise, telles que les fournisseurs ou les

prestataires. Ils doivent également respecter les mêmes critères que l'entreprise, ainsi que les normes de travaux internationales si elles sont situées à l'étranger par exemple.

Le **critère de gouvernance** est généralement le moins concret pour les épargnants. Il s'agit de l'ensemble des dispositifs assurant le bon fonctionnement de l'entreprise. Il comprend par exemple l'indépendance du conseil d'administration.

De manière générale, les critères ESG donnent des informations plus concrètes et précises des activités d'une société et de leurs impacts sur le monde autour. Ils sont de plus en plus pris en compte dans la création de nouveau fond par les sociétés de gestion, ou bien par les épargnants souscrivant directement à des actions. Ces critères seuls ne sont cependant pas suffisants pour évaluer la fiabilité financière d'une entreprise.

Choix des variables :

Dans le but de créer un modèle quantitatif pour voir la pertinence ou non de l'utilisation de ces facteurs dans la prédiction de performance financière, un choix de différentes variables a dû être effectué. Toutes les données proviennent de Bloomberg. Leurs sources sont donc assurées d'être fiables et précises au maximum.

Tout d'abord, il a fallu choisir des variables financières classiques. Elles sont obligatoires car elles sont la base des modèles. Il est impossible de prédire le prix d'une action future sans avoir le prix des actions passés par exemple. Pour ce faire, trois critères financiers ont été retenus : le prix de l'action (en euro), le revenu (en million) de l'entreprise et son Beta. (Le beta est la sensibilité du prix d'un actif par rapport à celle d'un marché de référence, ici le CAC40.)

Le prix de l'action est notre variable à prédire, synonyme de performance financière ou non de l'entreprise. Le revenu a été choisi car c'est une variable clé dans la santé d'une entreprise. Si la performance de celle-ci augmente, son revenu va la plupart du temps lui aussi augmenter et inversement. Le Beta quant à lui a été choisi pour traduire de la volatilité et faire le lien entre la tendance financière du moment avec celle de l'entreprise. S'il est élevé, une crise boursière impactera plus fortement le prix de l'action.

Il a fallu ensuite faire un choix de critère ESG pertinent. De manière générale, les données ESG ne sont pas répertoriées depuis de longues années pour la plupart (souvent depuis 2014), et de façon annuelle. Le même procédé sera ainsi appliqué : les données vont être transformées en format trimestriel.

Dans un premier temps, nous avons les « BESG score », qui se décomposent en quatre variables différentes : « BESG_ESG score », « BESG_environmental score », « BESG_social score », et « BESG_governance score ». Ce sont des scores attribués aux entreprises, d'une valeur de 1 à 10 (10 étant la meilleure note), qui évaluent leurs performances en matière de critère environnementaux, social, et de gouvernance. Le score global ESG est une agrégation des trois sous-scores mentionnés ci-dessus. Bloomberg utilise un mélange de données rapportées par les entreprises, des informations issues de tiers, et des algorithmes propriétaires pour calculer ces scores. Les étapes principales incluent :

- Collecte de données : Bloomberg rassemble des données ESG à partir des rapports d'entreprises, des déclarations réglementaires, des sites web d'entreprises, et d'autres sources publiques.
- Évaluation et pondération : Chaque critère est évalué et pondéré en fonction de son importance relative dans le secteur de l'entreprise. Les scores environnementaux, sociaux et de gouvernance sont calculés séparément.
- Agrégation : Les trois sous-scores sont ensuite agrégés pour former le score global ESG. Chaque sous-score peut être pondéré différemment selon les standards sectoriels et les meilleures pratiques de l'industrie.

Bloomberg veille à ce que ces scores soient transparents et reflètent fidèlement les performances ESG des entreprises, en mettant à jour régulièrement les données et les méthodologies utilisées.

Vient ensuite le « Disclosure ESG_score ». Celui-ci mesure le niveau de transparence des entreprises en matière de divulgation de leurs données et pratiques ESG. Ce score ne se concentre pas sur la performance des actions des entreprises, mais sur la quantité et la qualité des informations fournies. Ils se décomposent, comme pour le « ESG_score », en 3 sous catégories (environnemental, social, et gouvernance), et suivent la même méthodologie pour le calcul du score général. Ils sont notés de 1 à 100 (100 étant la meilleure note).

La sélection de ces deux catégories de variables est essentielle, car elles sont l'essence même du reflet des pratiques ESG générales des entreprises.

Enfin, les dernières variables ont été sélectionnées car elles donnent des informations sur un domaine en particulier. Les domaines choisis sont le résultat d'un choix personnel, effectué par ce que je pense être, les principaux problèmes dans le monde actuel quant à la production des entreprises. En effet, lorsque l'on entend les termes « production plus respectueuse », on pense

notamment au changement climatique, à la réduction de la dépense énergétique, ou encore à la gestion des déchets. C'est ainsi que les dernières variables sélectionnées sont :

- Si l'entreprise discute des opportunités d'investissement pour réduire le changement climatique. C'est une variable qui évalue si les entreprises abordent les opportunités créées par la transition vers une économie plus durable et à faible émission de carbone pour leurs nouveaux investissements. Elle prend comme valeur 1 si oui, 0 si non.
- SI l'entreprise met en place une politique de réduction de consommation énergétique, qui prend comme valeur 1 si oui, 0 si non.
- Le nombre de déchets rejeté, et le nombre de déchets recyclés, exprimés en tonnes métriques.

Les entreprises sont les principaux émetteurs de ces trois problèmes. En effet, la production de manière générale produit une forte quantité de gaz à effet de serre, à tous les niveaux, que cela soit pour la production (Industrie, construction), ou pour le transport (aviations, maritimes, ou par camions). Ces gaz sont les principaux responsables du réchauffement climatique.

L'Accord de Paris est un traité international sur le climat adopté par 196 parties lors COP21 à la Convention-cadre des Nations Unies sur les changements climatiques, tenue à Paris en décembre 2015. Il est entré en vigueur le 4 novembre 2016. L'objectif de cet accord et de renforcer et coordonner la réponse mondiale face au risque du réchauffement climatique, limitant l'augmentation de la température à 1,5 degrés. Pour respecter ces accords, il est essentiel que les entreprises cherchent à limiter leurs émissions de gaz à effet de serre.

Les entreprises se servent également de fortes quantités d'énergie pour leurs productions. La plupart de ces énergies sont les responsables de la production de Co₂, ce qui justifie la mise en place de politique d'économie, mais également car la quantité de la plupart de ces énergies n'est pas renouvelable.

« Une énergie non renouvelable est une source d'énergie qui se renouvelle moins vite qu'on ne la consomme et de manière négligeable à l'échelle humaine » [Énergie non renouvelable — Wikipédia \(wikipedia.org\)](https://fr.wikipedia.org/wiki/%C3%89nergie_non_renouvelable)

Selon l'Agence internationale de l'énergie (AIE), la part des énergies renouvelables dans la consommation d'énergie finale mondiale a augmenté, atteignant environ 20% en 2021. Cependant, cela indique qu'il reste tout de même une part importante de la production réalisée par des énergies qui ne le sont pas.

Enfin, l'émission de déchet non recyclé peut poser problème de par leurs quantités importantes, ce qui peut être synonyme de gaspillage ou d'une mauvaise manière de produire. De plus, certains déchets peuvent être toxique pour l'environnement. Les recycler permet donc, en plus de réduire le risque subi par la planète, de les réutiliser et in fine de réduire la quantité d'énergie et les matières premières utilisées.

Ces trois aspects, de par leurs importances, justifient leurs utilisations dans le modèle.

Ainsi, toutes ces variables seront retrouvées dans le modèle. Cependant, elles pourront subir des transformations pour une meilleur utilisation économétrique. Cela sera abordé dans la partie deux.

Choix des entreprises :

Pour le choix des entreprises, l'objectif a été d'en choisir avec des valeurs hétérogènes pour les variables expliquées plus haut. Le but étant de voir, quel que soit les pratiques ESG, quelles soit bonnes ou mauvaises, si l'implémentation de ces variables améliore la prédiction de la performance financière. Elles font bien évidemment toute partie du CAC40. Le choix s'est porté sur cinq entreprises :

- LVMH : ticker MC
- TOTAL : Ticker TTE
- LEGRAND : ticker LE
- RENAULT : ticker RNO
- VINCI : ticker DG

Legrand est Vinci possède un fort score de gouvernance avec une moyenne respectivement de 6.77 et de 6.1, tandis que LVMH a un score plus faible, avec une moyenne de 4.8

NB : Le score de transparence sur la gouvernance des entreprises est systématiquement haut, il semble que c'est un sujet sur lequel elles communiquent le plus, très certainement car elles sont obligées de fortement communiquer aux actionnaires leurs méthodes de gouvernance. De ce fait, il était impossible d'en trouver une avec un score faible dans ce domaine.

LVMH et Vinci possèdent un score environnemental et social plutôt faible (même si celui-ci s'améliore avec le temps), avec une moyenne de 1.77 et de 3.33 pour la première, et 2.28 et 0.75 pour Vinci, tandis que Total a un score plus élevé, avec une moyenne de 6.45 et de 5,25

La variable indiquant si l'entreprise discute des opportunités d'investissement est différente pour toute. Certaine entreprise, comme LVMH ou Legrand ne l'on jamais, d'autre comme Renault et Vinci le font depuis peu, et enfin Total le fait depuis plus longtemps.

Le prix de l'action est aussi très différent. LVMH a par exemple un prix bien plus haut que Renault et Vinci. Cela va permettre de voir si le modèle prédit aussi bien les prix élevés que faibles.

Enfin, les Betas de chacune d'entre différent. Total et Legrand ont un Beta historiquement défensif (inférieur à 1), tandis que celui de Renault est offensif (supérieur à 1). Pour LVMH et Vinci, celui varie, avec des passages où ils deviennent supérieurs à un, puis inférieur sur d'autres périodes.

B) Cadre légal, normes et labels ESG : Synthèse et lien avec la performance financière.

Les critères ESG ont pris de plus en plus d'ampleurs au cours de ces dernières années. En finance, cela entraîne l'apparition de nombreux cadres, labels et normes. Leurs rôles sont principalement de permettre une meilleure prise en compte des enjeux actuels pour les investisseurs, ce qui incite in fine les entreprises à adopter des pratiques plus respectueuses si elles veulent rester attractives. Des normes sont également mises en place pour obliger les entreprises à intégrer les critères ESG.

Voici les principales, nationales ou européennes, et qui ont donc un fort impact en France.

- Label ISR :

C'est un label français, créé en 2016 par le ministère des finances et de l'économie : Il vise à offrir aux épargnants une plus grande visibilité sur les produits financiers qui intègrent des critères environnementaux, sociaux et de gouvernance dans leur gestion. Il a donc un double objectif :

- Aider les épargnants à facilement choisir des investissements responsables.
- Encourager les fonds d'investissements à adopter une gestion plus responsable pour proposer ces fonds à leurs clients.

Pour qu'un fond/produit d'épargne obtienne ce label, il convient de respecter plusieurs critères et de suivre certaines étapes.

1- Vérifier que le fonds est éligible au label :

Un audit est réalisé par un des 3 organismes accrédités (Afnor certification, Deloitte, EY France). Ils vérifient si le fond est éligible.

« De manière générale, les sociétés de gestion des fonds labellisés doivent se fixer des objectifs ESG (environnement, social et gouvernance), mettre en place une méthodologie et des moyens d'analyse suffisants et prendre en compte les résultats de ces analyses dans la construction et la gestion du portefeuille. Enfin, elles doivent rendre compte de l'atteinte des objectifs fixés auprès des investisseurs notamment par le calcul d'indicateurs mesurables. »

2- Vérifier que le fonds remplit les critères de labellisation

Le fond doit respecter les critères du cahier des charges :

- Les objectifs généraux (financiers et ESG) sont précisément définis et décrits aux investisseurs et pris en compte dans la politique d'investissement du fonds.
- Définition de l'analyse et la notation des critères ESG mise en œuvre par les entreprises dans lesquelles le fonds investit
- La prise en compte des critères ESG dans la construction et la vie du portefeuille
- La politique d'engagement ESG avec les entreprises dans lesquelles le fonds investit
- La transparence de gestion du fonds
- La mesure des impacts positifs de la gestion ESG sur le développement d'une économie durable.

Une fois le rapport d'audit terminé, l'organisme (indépendant) du label attribue ou non la certification pour une durée de trois ans renouvelables. Cependant, des contrôles continuent d'être effectués durant cette période.

- **Norme CSRD** ou Corporate Sustainability Reporting Directive

La CSRD, directive adoptée en 2022 et qui succède à la NFRD (Non Financial Reporting Directive), est une directive européenne qui s'applique à toutes les grandes entreprises qui dépassent deux des trois seuils suivants : 250 employés, 40 millions d'euros de chiffre d'affaires, 20 millions d'euros de total de bilan, et à toutes les entreprises cotées en bourse dans l'UE, à l'exception des micro-entreprises. Elle vise à améliorer la transparence des informations ESG des entreprises, avec des rapports standardisés pour une meilleure comparaison. Elles doivent publier comment elles intègrent ces critères. Les informations doivent inclure des descriptions de la gouvernance et des processus de gestion des risques, des impacts des activités sur les questions ESG, et des objectifs et mesures de performance ESG. Ces rapports sont ensuite audités par un tiers indépendant.

Cette norme permet ainsi aux investisseurs des prises de décisions plus responsables grâce à une meilleure information des pratiques des entreprises, et celle-ci subissent une plus forte pression quant à leur véritable mise en place de meilleure pratique. De plus étant une norme européenne, les rapports standardisés impliquent une harmonisation des normes de l'union.

- **Taxonomie européenne :**

Créée en 2020, La taxonomie européenne est un cadre de classification conçu par l'Union européenne pour définir ce qui constitue une activité économique durable sur le plan environnemental.

La taxonomie vise à fournir une définition claire et standardisée des activités économiques considérées comme durables. Elle a pour but de diriger les flux de capitaux vers des activités qui soutiennent les objectifs environnementaux de l'UE, contribuant ainsi à la réalisation des objectifs climatiques et énergétiques de l'Union. C'est son cœur de stratégie.

La taxonomie identifie six objectifs environnementaux principaux :

- Atténuation du changement climatique
- Adaptation au changement climatique
- Utilisation durable et protection des ressources aquatiques et marines
- Transition vers une économie circulaire
- Prévention et contrôle de la pollution
- Protection et restauration de la biodiversité et des écosystèmes

Un de ses principes fort et également le « Do No Significant Harm" (DNSH) : Une activité économique doit contribuer à un des 6 critères, sans porter de préjudice aux autres. Les activités doivent également être socialement responsable.

Elle oblige aussi à une transparence des informations et un reporting standardisé, qui, couplé à la CSRD permet de meilleure information pour les investisseurs. Les entreprises doivent se plier à la taxonomie (sous menace de sanction telle que des amendes), indiquer la proportion de leurs activités économiques alignées avec les critères de la taxonomie, puis fournir un reporting détaillé de comment elles se conforment à ces critères dans le cadre de la CSRD.

Les greens bonds : leurs importances dans l'ESG et leurs liens avec la taxonomie européenne.

Les green bonds sont des obligations émises pour financer des projets environnementaux spécifiques tels que les énergies renouvelables, l'efficacité énergétique, la gestion des déchets, la conservation de l'eau, et d'autres initiatives écologiques. Les green bonds peuvent utiliser les critères techniques de la taxonomie pour définir quels projets sont éligibles au financement. Par

exemple, un projet de production d'énergie solaire doit respecter des critères spécifiques d'efficacité et de durabilité définis par la taxonomie.

L'utilisation de la taxonomie permet une harmonisation des critères des green bonds. Ils sont donc comparables entre eux, et permettent une confiance accrue des investisseurs qui peuvent voir précisément lesquelles sont les plus intéressants.

Le greenwashing : une limite importante de l'ESG :

Un des aspects importants non évoqué jusqu'à présent est celui du Green Washing.

« Le greenwashing est un procédé de marketing ou de relations publiques utilisé par une organisation pour se donner une image trompeuse de responsabilité écologique. La plupart du temps, les dépenses consenties concernent davantage la publicité que de réelles actions en faveur de l'environnement et du développement durable. » [Greenwashing — Wikipédia \(wikipedia.org\)](https://fr.wikipedia.org/wiki/Greenwashing)

C'est une des limites principales à l'inefficacité des critères ESG. Ce procédé réduit la confiance de la population quant aux réelles actions, car il devient difficile de faire la différence entre des entreprises réalisant du Greenwashing et celle appliquant réellement les critères ESG. De plus, les dépenses normalement utilisées pour produire plus respectueusement sont alors allouées au marketing.

Les normes CSRD, et la taxonomie européenne, qui visent à améliorer la transparence, permettent ainsi d'efficacement lutter contre ce phénomène. Les reportings détaillés et standardisés, qui obligent par exemple à dévoiler les dépenses d'investissement dédiées à l'ESG, indiquent aux investisseurs leurs véritables actions et politique environnemental, social et de gouvernance.

Pourquoi les critères ESG permettent-ils une amélioration de la performance financière ?

Malgré les enjeux environnementaux actuels, il faut garder à l'esprit que la performance financière reste la préoccupation première de nombreux investisseurs. Ces normes, labels et autres... ne doivent pas, au minimum, réduire la performance globale des marchés financiers.

Tout d'abord, lesdites créations de normes citées précédemment sont synonymes de sanction, souvent financière, en cas de non-respect. Il est donc évident que les respecter permet de ne pas subir de perte liée à des actions légales.

Ensuite, les critères ESG permettent une réduction de plusieurs risques, ce qui est in fine, synonyme de performance.

Risque climatique :

« Le risque climatique est un risque lié à la vulnérabilité accrue des entreprises par rapport aux variations des indices climatiques (température, précipitations, vent, neige...) ».

Dans de nombreux secteurs économiques agriculture et agroalimentaire, énergie, textile, tourisme, loisirs, construction..., le risque climatique est beaucoup plus important que le risque de marché traditionnel (taux d'intérêt, change, matières premières...) et peut expliquer à lui seul l'essentiel de la volatilité du résultat. » [Risque climatique — Wikipédia \(wikipedia.org\)](https://fr.wikipedia.org/wiki/Risque_climatique)

Une entreprise soucieuse de ces potentiels problèmes sera mieux préparé pour y répondre, et un ensemble d'entreprise soucieuse permettra ainsi de réduire le risque pour l'ensemble du secteur. Par exemple, une production d'un secteur entier de l'agroalimentaire respectueuse des sols, va permettre de réduire le risque global de l'appauvrissement de la terre pour la production de cette denrée. Les performances financières seront ainsi supérieures.

Les risques sociaux :

Le risque social fait référence aux menaces ou aux impacts négatifs potentiels que les actions ou les inactions d'une entreprise peuvent avoir sur les parties prenantes internes et externes, y compris les employés, les communautés locales, les clients, les fournisseurs et la société en général. Ces risques peuvent affecter la réputation, les opérations, les performances financières et la durabilité à long terme de l'entreprise.

Une entreprise qui prend des mesures pour ses employés, son approvisionnement de matière première ou autre, voit sa probabilité de connaître des événements tels que des grèves, ou une mauvaise image de marque, diminuer. Par exemple, certaines marques de « Fast Fashion », tel que Shein, connaissent une forte baisse de popularité suite au scandale des Ouïghours.

Enfin, les critères ESG accélèrent l'émergence et le développement de marché lié à l'environnement ou au social. La demande croissante de produit respectueux de la planète, ou de technologie verte, en fait un marché rentable et intéressant pour les entreprises. De plus, ces

règles et normes permettent aux entreprises d'accéder plus facilement à des financements comme vu précédemment (green bonds par exemple), ce qui accélère leurs croissances et donc leurs performances.

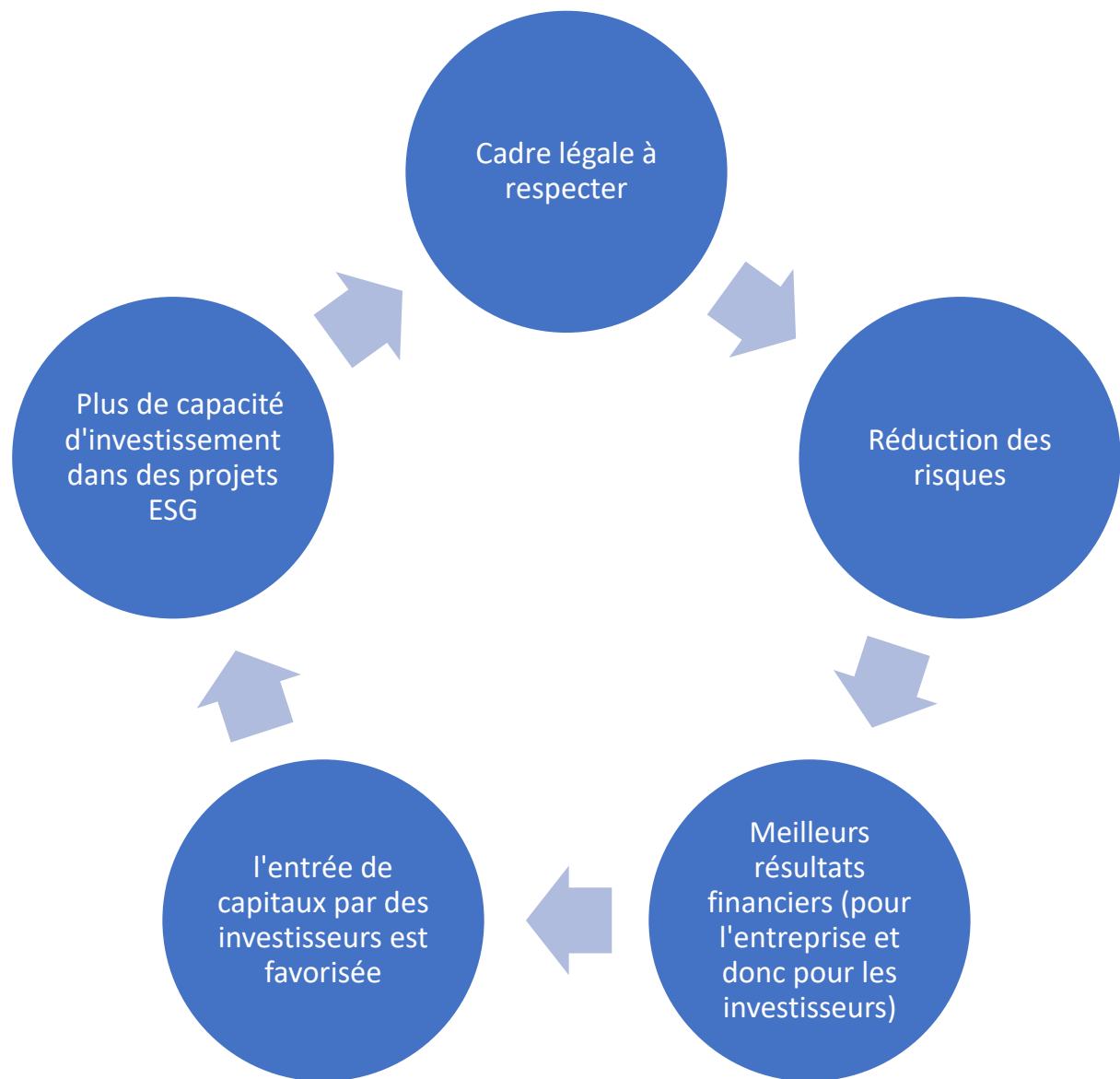
Conclusion entre le développement d'un cadre ESG et la performance financière.

Comme observé, une production respectant les critères ESG sera de facto plus efficace à long terme grâce à notamment, une réduction de certains risques, ou de nouvelles opportunités de marché. Cependant, le développement d'un cadre ESG global a un double effet :

- Créer un cadre légal, pour obliger les entreprises à respecter des normes et critères sous peine de sanction.
- Inciter les investisseurs (et donc la population) à investir dans des projets plus respectueux socialement et de l'environnement

Dans le monde actuel, l'action coordonnée de population semble au moins aussi efficace que le pouvoir des instructions sur les plus grandes entreprises, si ce n'est plus. La détérioration de l'image d'une entreprise peut souvent mener à sa perte, et le boycott est une arme de plus en plus utilisée par exemple.

Le but est donc de créer un cercle vertueux pour les entreprises respectant les critères ESG. Les agents vont de plus en plus investir dans ces entreprises, que cela soit directement en achetant des actions et obligations de celles respectant la taxonomie, soit au travers de produits d'épargne ou de fonds d'investissements, qui respectent le label ISR. Ceci augmente le capital de ces entreprises, qui peuvent donc de nouveau investir dans des projets durables et de facto avoir un meilleur résultat, ce qui renforce la performance financière des actifs achetés par les agents.



Maintenant que les bases théoriques ont été posé et expliqué, il est possible de commencer la partie quantitative de l'analyse, avec l'analyse portée sur l'intégration des variables.

2 – L’analyse quantitative : Implémentation des variables et création du modèle.

A) Intégration des variables dans le modèle.

La sélection et l’implémentation des variables sont des étapes cruciales dans la création de modèles quantitatifs. Elles influencent directement la qualité, la précision et l’interprétabilité du modèle.

Dans la partie précédente, nous avons vu les variables sélectionnées et pourquoi elles l’ont été. Cependant, l’utilisation de variable dans un modèle nécessite une réflexion et une analyse préalable. Des modifications peuvent y être apportées, par exemple pour régler des problèmes de temporalités, ou en cas de données incomplètes (NA). De plus, le lien entre certaines variables peut ne pas correctement refléter les pratiques ESG des entreprises, ou le modèle ne pourrait pas les capter avec précision.

Dans la pratique commune actuelle, il est courant d’implémenter directement les scores ESG globaux dans les modèles. Ces scores sont souvent utilisés comme variables explicatives pour évaluer leur impact sur les rendements financiers. Les autres variables, telles que celles liées aux déchets ou à la production de CO₂ sont quant à elles moins utilisées dans ce genre de modèle, mais plus souvent dans le but de prédire les couts potentiels liés aux politiques de transition énergétique, ou voir l’efficacité des actions de recyclage pratiquées par l’entreprise. Ces indicateurs sont donc moins présents dans des modèles purement financiers général, mais utilisé pour des cas plus précis. Il sera donc intéressant de voir si leur présence dans ce modèle est pertinente.

Dans notre modèle, l’objectif est de prédire le prix futur des actions des entreprises. La variable à expliquer est donc le prix de l’action, tandis que les variables explicatives seront le Beta et le revenu pour le modèle uniquement financier, et les variables y seront rajoutées ensuite pour créer le modèle complet.

Tout d’abord, le revenu des entreprises n’est calculé que par trimestre, et non de façon journalière comme pourrait être le prix de l’action ou le Beta. Cela peut créer des incohérences dans la façon dont ces variables sont alignées dans le temps et entraîner des difficultés dans

l'association des données et des erreurs dans le modèle. Plusieurs solutions étaient possibles pour remédier à ce problème, mais celle d'exprimer le prix de l'action et la Beta en trimestre a été retenu au vu de la temporalité des autres variables. Pour ce faire, une moyenne du prix et du Beta sera calculée sur les trois mois. Le résultat du modèle sera alors moins précis, car il ne pourra prédire qu'une moyenne du prix de l'action des trois prochains mois, mais il devrait cependant toujours permettre de dégager une tendance dans la performance financière futur des entreprises.

Le même problème se pose pour toutes les variables ESG car elles sont calculées de façon annuelle. Ainsi, elles seront transformées pour être exprimées en format trimestriel également. Pour ce faire, la méthode utilisée sera plus complexe qu'un simple calcul de moyenne. En effet, un score ESG d'une entreprise est exprimé à la fin de chaque année, le 31/12. Cependant, ses pratiques s'améliorent (ou se détériorent) tout au long de l'année. De ce fait, il est possible de poser l'hypothèse qu'un score ESG évolue tout au long de l'année d'une manière plus ou moins constante. Ainsi, les variables subiront une hausse linéaire sur chaque trimestre, ce qui retranscrira au mieux cette augmentation, en suivant cette méthode de calcul.

$$\text{Score_ESG } T - \text{Score_ESG } T-1 = x$$

$x/4$ = Hausse à appliquer à chaque trimestre.

Par exemple, si une entreprise a un score ESG de 2,5 en 2020, et de 2,6 en 2021, elle aura 2,525 au 1^{er} trimestre, 2,55 au deuxième, 2,575 au troisième, et 2,6 pour le quatrième.

Une autre réflexion à avoir est celle du lien entre les variables exprimant le score ESG, environnemental, social et de gouvernance, et le score de transparence associée, en fonction des documents et rapports produit par l'entreprise, qui permettent leurs calculs. Le score est calculé en fonction des actions prises par les entreprises, cependant ce score ne prend donc pas en compte le niveau de transparence. Une entreprise peut donc avoir théoriquement un très bon score en ne publiant que les documents pour lesquels ils sont obligés.

Or, est-il possible de correctement interpréter les pratiques ESG d'une entreprise si les informations publiées ne sont pas nombreuses. Cette mauvaise transparence peut donc poser des problèmes de pertinence des autres variables ESG et donc fausser notre modèle. Pour le résoudre, aucune des deux variables du score ou de la transparence ne sera utilisé, mais une troisième sera créée, résultant de la multiplication des deux précédentes. Cette nouvelle

variable, appelée « ESG_score » aura ainsi une note sur 1000. Elle permet de capturer l'interaction entre score ESG et transparence. Cela signifie qu'une entreprise avec un score ESG élevé mais une faible transparence aura un score global plus bas. De plus, la combinaison des deux permet de réduire le nombre global de variable dans le modèle, ce qui peut aider à éviter un problème de multi-colinéarité. Cette multiplication est cependant seulement possible si on pose l'hypothèse que les deux variables ont la même importance relative au départ. Sachant que notre analyse se porte sur la pertinence des critères ESG en général dans un but de prédiction, cette hypothèse semble être raisonnable.

Ensuite, il faut s'occuper des variables liées aux déchets. Les données brutes n'expriment qu'une quantité de déchets émis et recyclés. Ces quantités dépendent de la production totale d'une entreprise. Au plus une entreprise produite, au plus elle aura de déchet. De plus, une qui ne produit que quinze tonnes métriques de déchets mais n'en recycle qu'une, ne semble pas avoir de meilleur pratique qu'une autre qui produit vingt tonnes mais en recycle dix. Ce qui nous intéresse ici, ce n'est pas réellement la production de déchets, mais la proportion de ceux qui sont recyclés.

Ces deux variables ne seront donc pas utilisées mais combinées, pour créer un ratio de déchets recyclés, nommé « Waste_ratio ». Il va permettre de bien capter la relation entre les deux variables, tout en gardant les informations qui nous intéressent dans chacune d'entre elles. Il mettra bien en lumière si les pratiques de gestion des déchets s'améliorent ou non au fil des années. Il sera exprimé en valeur décimale et non en pourcentage, pour ne pas prendre le risque de poser un quelconque problème au modèle.

Exemple : en 2015, LVMH a produit 92,826 tonnes métriques de déchet, pour 27,848 tonnes recyclées. On a donc :

$$27,848/92,826 = 0,30.$$

LVMH a recyclé 30% de ses déchets en 2015. Ce résultat est bien plus révélateur des pratiques de l'entreprise que simplement les 27,848 tonnes que l'on a au départ.

Enfin, il reste le cas des valeurs manquantes (NA) et des titres à traiter. Les entreprises étudiées ayant été choisies spécifiquement pour les données sélectionnées, il y a donc très peu de NA. Seul Total n'a pas renseigné sa quantité de déchets émis et recyclés pour les années 2016 et 2017. Les quantités ne semblent néanmoins ne pas beaucoup varier d'une année à l'autre pour cette entreprise, les NA seront donc remplacées par la médiane, dans l'objectif de garder une

cohérence. Avec ce procédé, la forme de distribution générale des données n'est pas modifiée. De plus, l'échantillon étant relativement faible, choisir la médiane plutôt que la moyenne est pertinente car celle-ci est moins sensible aux fluctuations. De même, il est difficile de modéliser les valeurs manquantes de par la petite taille de l'échantillon.

Cette méthode crée cependant une uniformité artificielle des valeurs. D'autres variables ayant déjà des valeurs identiques, cela ne devrait pas trop influencer négativement sur le modèle. Par exemple, les données liées à la gouvernance sont souvent identiques d'année en année, les entreprises ne modifiant pas leurs méthodes de gestion régulièrement. Il y aura néanmoins forcément une légère perte d'information, car la médiane ne prend pas en compte les relations entre les données manquantes et d'autres variables.

Les titres, proviennent de Bloomberg, sont eux à modifier pour éviter tout problème lors de la programmation. En effet, certains caractères ou espaces peuvent mal être lu par le logiciel et des titres trop longs ou trop compliqués peuvent tout simplement faire perdre du temps car difficiles à écrire.

Par exemple, les scores ESG sont renommées simplement « ESG_score, Environmental_score, Social_score et Gouvernance_score », ou « climate change opportunity discussion » devient simplement « Opportunity » et « Energy reduction initiative » devient « Energy_reduc »

Il faut noter que les variables « climate change opportunity discussion » et « Energy reduction initiative » sont des variables binaires. Elles ne subiront aucune modification, étant utilisables comme t'elles.

Période d'analyse choisie :

La dernière interrogation qu'il reste pour créer notre modèle est la période d'analyse. Les critères ESG ne sont apparues que récemment dans les préoccupations mondiales, et la collecte de données n'y fait donc pas exception. De ce fait, les premiers scores ESG à être répertoriés remontent à 2015. De plus, ils sont calculés dans le courant de l'année N+1 pour l'année N. De ce fait, les scores de 2023, calculés normalement en 2024, ne sont pas encore disponible en juin. Il est, de ce fait, impossible de prédire les prix de 2023 pour tester notre modèle dû au manque de valeurs pour cette année. La période d'analyse se portera donc sur la période 2015-2022, avec comme but la prédiction les prix de 2022, période imposée par le faible nombre de données disponible. Prédire les prix de 2022 permettra ainsi de tester la précision du modèle sur des prix connus.

Il est possible d'effectuer toutes ces transformations directement sur Rstudio, le logiciel utilisé pour coder. Cependant, pour simplifier la compréhension du code, les modifications seront faites directement sur Microsoft Excel à l'aide des formules disponibles. La base de données sera aussi divisée en deux : Une base d'entraînement « Nom_de_l'entreprise_Train », utilisée pour entraîner notre modèle. Celle-ci comprend ainsi les données de 2015 à 2021. Puis, une base de test sera créée, « Nom_de_l'entreprise_Test », comprenant les prix des actions de 2022, pour voir si notre modèle permet de correctement prédire le prix.

Chaque entreprise aura ses deux propres bases de données d'entraînement et de test. Là encore, le but est de simplifier la compréhension du code.

Ce processus de transformation prend du temps en plus d'être fastidieux car la moindre erreur peut fausser l'entièreté du modèle et donc de l'analyse. Une fois ces étapes réalisées, il est possible de passer à la création du modèle.

B) Création du modèle et analyse des résultats.

En finance quantitative, et plus généralement en Data science, il existe de nombreux modèles, chacun ayant leurs avantages et inconvénients. Le choix de celui à utiliser est donc issu d'une analyse et d'une réflexion, souvent réalisée grâce à une étude statistique de la base de données au préalable. Cette étape est cruciale et ne doit pas être négligée. Une fois faite, un choix d'un ou plusieurs modèles sera/seront choisi, puis la précision des résultats sera testée pour choisir celui donnant les meilleurs résultats.

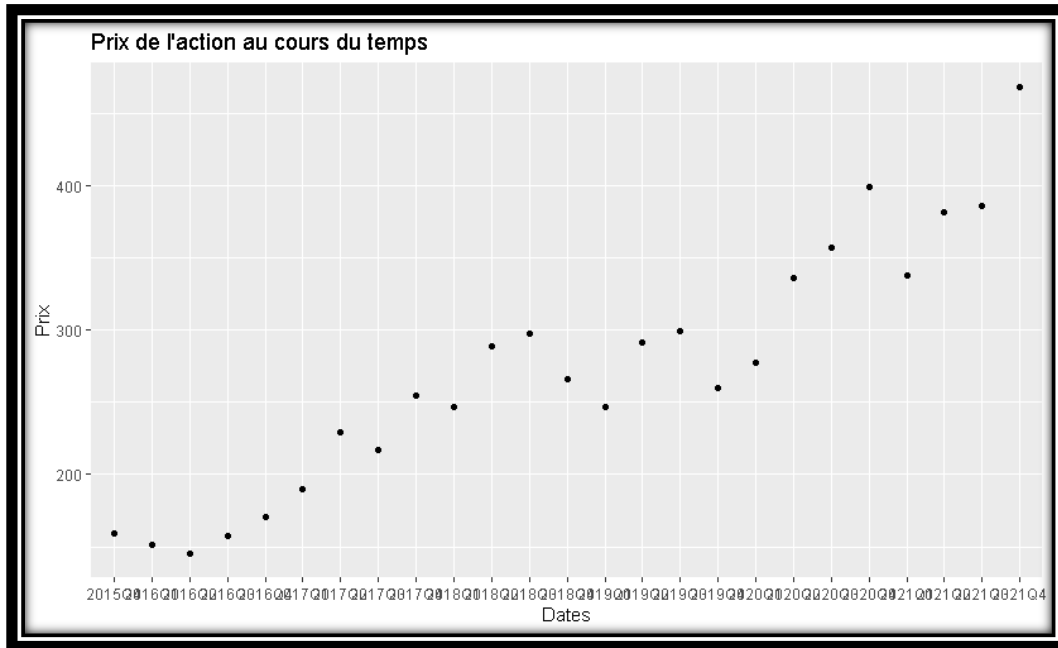
Il faut cependant noter qu'un modèle trop précis n'est pas obligatoirement le meilleur : Dans certaines situations, le besoin d'interpréter les coefficients des variables est plus important que le résultat en lui-même. Dans la plupart des cas, un modèle plus précis est difficile à interpréter. Dans le cas de ce travail, ce n'est pas le besoin premier. Cela nous donne donc accès à davantage de meilleurs modèles pour notre prédiction. De plus, un résultat trop précis peut résulter d'un sur-apprentissage. Le sur-apprentissage est un phénomène qui se produit lorsqu'un modèle économétrique est trop bien ajusté aux données d'entraînement. Cela signifie que le modèle capture non seulement les tendances générales présentes dans les données, mais aussi le bruit et les fluctuations aléatoires. En conséquence, même si le modèle montre une très bonne performance sur les données d'entraînement, il aura une faible performance sur des données nouvelles ou non vues, car il ne généralise pas bien. Là encore, ce problème ne concerne pas vraiment cette analyse. Même s'il faut évidemment éviter le plus possible ce phénomène, le modèle n'ayant pas vocation à être utilisé sur de nouvelles données, mais seulement à voir s'il est plus précis avec des variables ESG.

Enfin, cette analyse sert normalement à sélectionner les données les plus pertinentes à utiliser dans le modèle, et détecter des corrélations trop importantes. Cependant, la problématique de ce travail vise à voir la pertinence des variables ESG de façon générale, et non à savoir quelles sont les meilleurs. En enlever pourrait ainsi ne pas être pertinent, donnant trop de poids aux variables financières présentes. Le phénomène de corrélation sera lui abordé plus tard dans l'analyse.

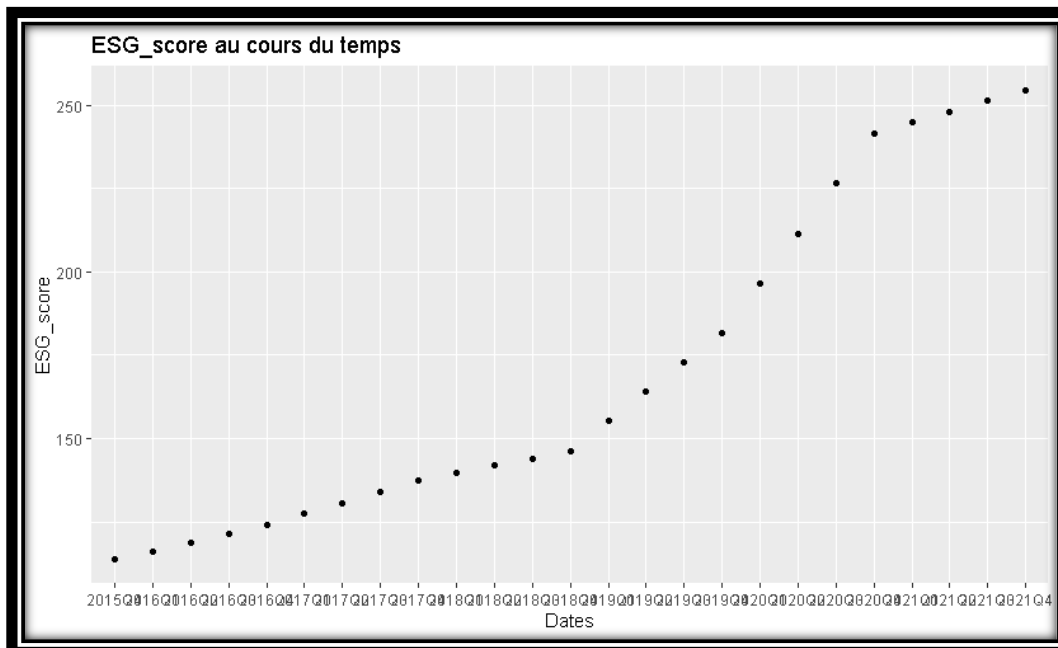
Etude statistique :

Tout d'abord, une analyse graphique de la variable cible (le prix) et de la variable cible principale (le score ESG) ont été réalisés pour LVMH.

***NB :** Seul l'étude statistique de LVMH sera présentée ici pour éviter la redondance. En effet, la méthode d'analyse est la même quelque que soit l'entreprise. La tendance émergente des graphiques et les analyses sont presque identiques pour toutes.*



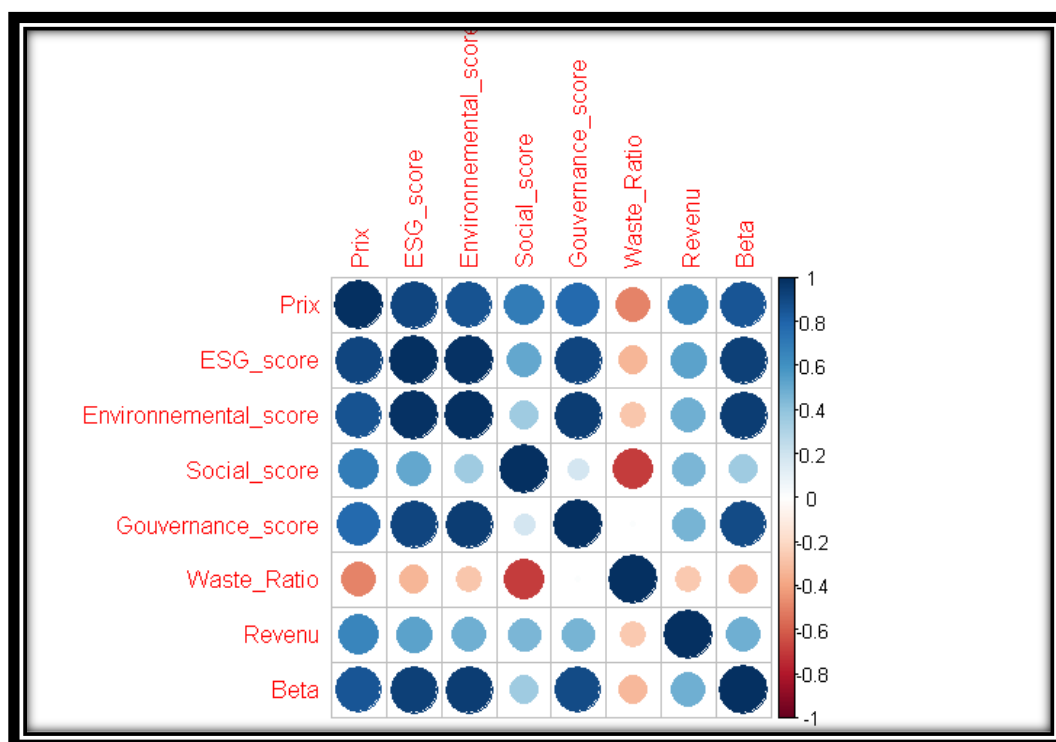
```
graph1 <- ggplot(Train, aes(x = Dates, y = Prix)) + geom_point() + labs(title = "Prix de l'action au cours du temps")
```



```
graph2 <- ggplot(Train, aes(x = Dates, y = ESG_score)) + geom_point() + labs(title = "ESG_score au cours du temps")
```


Ces représentations permettent dans un premier temps d'analyser la tendance générale des données. La moyenne trimestrielle du prix de l'action augmente au cours du temps malgré quelque légère baisse, tandis que le score ESG de l'entreprise ne fait que s'améliorer durant la période. Il ne semble pas y avoir de valeurs aberrantes, et la période d'analyse est trop faible pour voir apparaître une saisonnalité ou des cycles de manière frappante ainsi que des ruptures. Il pourrait y avoir cependant une légère corrélation.

Il faut donc tester cette corrélation entre ces deux variables, et même de façon générale entre toute. Pour ce faire, une matrice de corrélation a été réalisé avec l'entièreté des variables, excepté les variables binaires, avec lesquels la corrélation ne peut être calculé du fait de leurs valeurs qui est 1 ou 0.



```
correlation_matrix <- cor(Train[, c("Prix", "ESG_score", "Environnemental_score", "Social_score", "Gouvernance_score",
"Waste_Ratio", "Revenu", "Beta")))
```

```
Matricecor <- corrplot(correlation_matrix, method = "circle")
```

Cette matrice fait ainsi apparaître une forte relation entre le score ESG général, et le score environnemental, social et de gouvernance. Cela semble logique, car même si le score ESG global n'est pas une simple moyenne des autres scores comme expliqué précédemment, il n'en reste pas moins réalisé à l'aide de ceux-ci. De plus, le prix et le Beta semblent eux aussi connaître une relation importante avec ces scores. Même si celle-ci peut nous desservir dans la création de notre modèle, elle renforce néanmoins l'étude précédente du lien entre les pratiques ESG et l'amélioration de la performance financière de l'entreprise.

Seule la variable « Waste_Ratio » semble avoir une relation négative avec les autres variables. Il est difficile de savoir exactement pourquoi, mais il se pourrait que cela vienne du faible ratio de déchet recyclé que LVMH possède.

Au vu de ces corrélations, et du choix de ne pas supprimer de variable, il est désormais primordial de choisir un modèle étant robuste à la corrélation.

Le non-choix de la régression linéaire multiple :

Un des premiers modèle auquel on pense de par sa simplicité est la régression linéaire. Elle permet notamment d'interpréter les coefficients avec précisions, ce qui donne l'impact de chaque variable explicative sur la variable cible. De plus, elle fonctionne correctement sur des ensembles de données de taille modérés comme ici. Comme expliqué au-dessus, même si interpréter les coefficients n'est pas le but premier de ce travail, il est toujours intéressant de voir l'impact des variables ainsi que leurs pertinences dans le modèle dans le cadre d'une analyse futur, car elle permet bien souvent de faire apparaître les variables les plus utiles. Cependant, l'analyse statistique précédente et la façon dont la temporalité des variables a été traité rendent impossible son utilisation. En effet, cette méthode n'est absolument pas robuste à la colinéarité entre les variables. Lorsque les variables indépendantes sont colinéaires, les coefficients estimés par la régression linéaire deviennent très sensibles aux petites modifications des données. Cette instabilité se manifeste par des coefficients de régression qui peuvent avoir des valeurs très grandes et des signes incohérents. En conséquence, les prédictions deviennent moins fiables et les coefficients ininterprétables. De plus, les variables ESG annuels ont été transformées en variable trimestrielle en posant l'hypothèse que l'amélioration des pratiques ESG au cours d'une année se fait d'une façon linéaire. Or, la régression linéaire, comme son nom l'indique, suppose une relation linéaire entre la variable

cible et les autres. Créer artificiellement de la linéarité au sein des données comme ici peut ainsi fausser le modèle.

Le choix du modèle :

Il existe plusieurs modèles étant robuste à la corrélation, et ne se basant pas sur une relation linéaire entre les variables. Ainsi, (et au vu de mes compétences économétriques), le choix s'est porté sur deux modèles :

- Un relativement simple, le random forest
- Un plus compliqué, le XGboost,

Le random forest est un algorithme d'apprentissage automatique basé sur le concept d'ensemble. Il combine les prédictions de plusieurs arbres de décision pour obtenir une prédiction plus robuste et précise. La prédiction finale est obtenue par un vote majoritaire parmi tous les arbres. Il est moins affecté par la colinéarité ou les valeurs aberrantes et capture les relations complexes entre variables. Il peut être cependant plus lent à entraîner, et nécessiter une plus grande base de données pour éviter un surajustement, ce qui peut être un problème ici.

Le XGBoost est un algorithme de boosting. Il construit des arbres de décisions en améliorant les erreurs des modèles précédents en utilisant une méthode appelée « gradient boosting ». De plus, il intègre une régularisation pour éviter le surajustement, ce qui améliore la généralisation du modèle. Il a un haut niveau de précision.

Il capte également les relations complexes, et peut intégrer une grande variété de données. Il peut gérer efficacement les données manquantes, ce qui peut être utile ici pour Total. Il est cependant plus gourmand en ressource et peut être long à entraîner sur des grands ensembles de données.

Dans les deux cas, les résultats des coefficients sont difficiles (voir impossible) à interpréter. Il sera donc impossible de détecter l'impact précisément des variables sur le prix. Ces deux modèles seront utilisés pour voir lequel fournit les meilleurs résultats, mais également pour éviter des incohérences. Même dans le cas où l'un est moins précis que l'autre à l'issue de l'étude, les deux devront montrer une conclusion similaire. Si l'un indique que les critères ESG

ne sont pas pertinent à modéliser tandis que l'autre si, il y aura de grande chance qu'une erreur ai été commise.

Cas de LVMH :

Modèle financier :

Pour créer ce modèle, les deux variables financières, les revenus et le Beta sont utilisées.

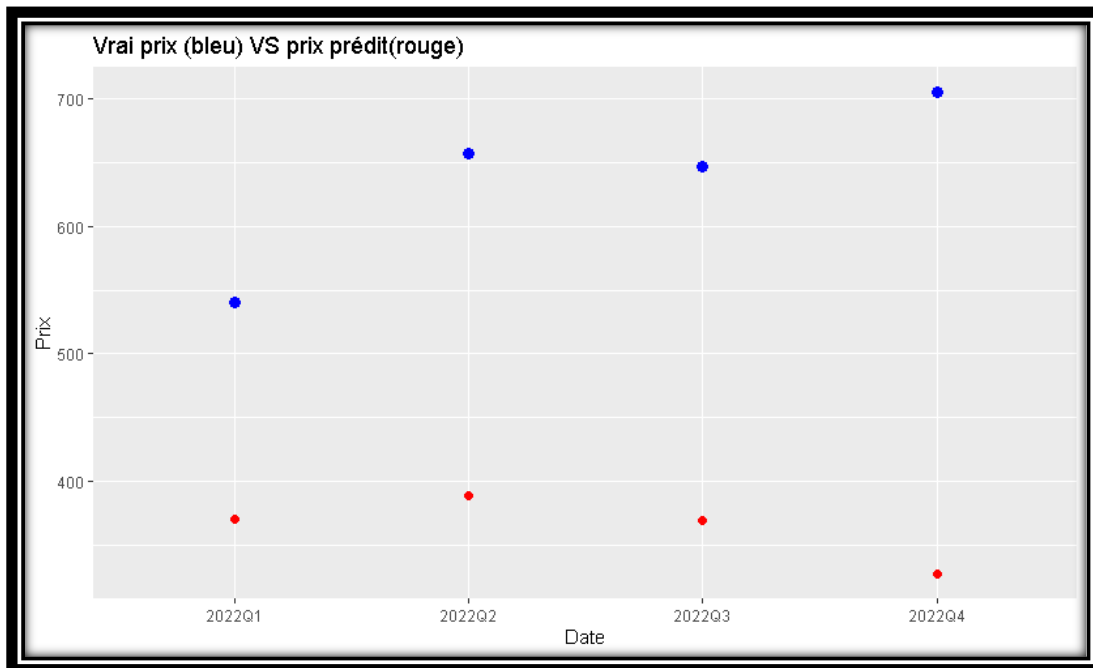
Lorsque l'on applique le random forest sur la base d'entraînement, il affiche une moyenne des résidus carrés (MSE) est de 2196.163. Au plus ce résultat est bas, au mieux le modèle est, cependant ce chiffre en lui-même doit être comparé avec les autres MSE pour pouvoir l'utiliser. Le pourcentage de la variance expliquée est de 69,17%. Le revenu et le Beta expliquent environ 69% du prix de l'action. Au plus ce chiffre est élevé, au plus le modèle est performant, mais là encore il convient de le comparer avec les résultats des autres modèles.

Pour tester la véritable capacité prédictive du modèle, il faut l'appliquer ensuite sur la base de teste. La précision de celui est mesurée grâce à la racine carrée du MSE (RMSE) est du coefficient de détermination (R2). Le RMSE indique la racine carrée de la moyenne des erreurs entre les valeurs prédites et les valeurs réelles du prix de l'action. Ici, sa valeur est de 283.10, ce qui est très élevé. Les prédictions du modèle sont en moyenne loin des valeurs réelles. Le R2 indique lui le pourcentage de la variance expliqué par le modèle. Il est quant à lui négatif, ce qui n'est pas normal (il est normalement compris entre 0 et 1). Cela indique que le modèle est pire qu'un modèle de base qui prédit simplement la moyenne des valeurs de la variable cible.

Résultat de la prédiction :

Dates	Vrai prix	Prix prédit
2022-T1	540.7	370.68
2022-T2	657.1	389.21
2022-T3	646.8	369.48
2022-T4	705.4	327.22

Si on représente graphiquement la prédiction, on obtient :



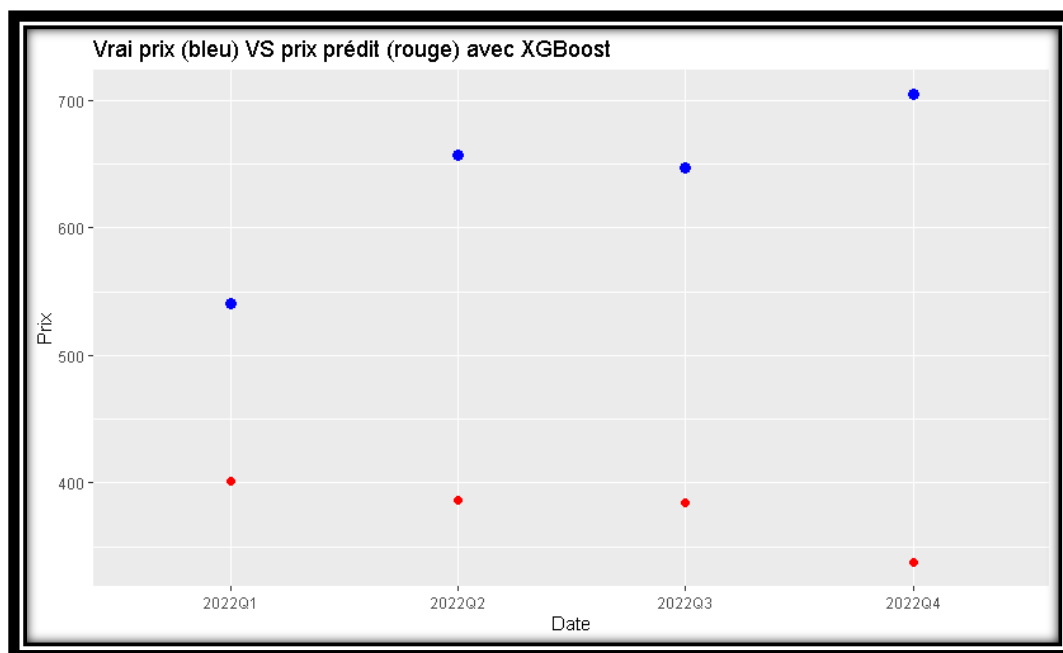
```
ggplot(graph_data, aes(x = Dates, y = Actual)) + geom_point(color = "blue", size = 3) + geom_point(aes(y = Predicted), color = "red", shape = 21, fill = "red", size = 2) + labs(title = "Vrai prix (bleu) VS prix prédit(rouge)", x = "Date", y = "Prix")
```

Comme on peut le voir, les résultats sont mauvais. Plusieurs raisons sont possibles : Le modèle a pu sur-apprendre ou sous apprendre. Il y a peu de chance que cela soit du a la qualité des données en elle-même, les valeurs de l'année 2022 étant relativement proche de celle de 2021. Il n'y a pas eu de grand changement sur le score ESG par exemple.

Pour le XGboost, a l'instar du random forest, il n'est pas possible d'analyser les performances du modèle sur la base d'entraînement. Il faut directement l'appliquer sur la base de test et voir la précision du modèle. On applique les mêmes mesures de précision et on obtient un RMSE de 272,64 et un R2 de -19,57, ce qui est encore synonyme de mauvaise prédiction.

Les résultats sont :

Dates	Vrai prix	Prix prédit
2022-T1	540.7	401.72
2022-T2	657.1	386.17
2022-T3	646.8	384.09
2022-T4	705.4	337.14



```
ggplot(graph_data3, aes(x = Dates, y = Actual)) + geom_point(color = "blue", size = 3) + geom_point(aes(y = Predicted),
color = "red", shape = 21, fill = "red", size = 2) + labs(title = "Vrai prix (bleu) VS prix prédit (rouge) avec XGBoost", x =
"Date", y = "Prix")
```

Comme on peut le voir, les résultats restent mauvais, malgré une légère amélioration. Bien que meilleur, le XGboost ne parvient tout de même pas à capter assez de relation entre les variables pour correctement prédire les valeurs.

Modèle complet :

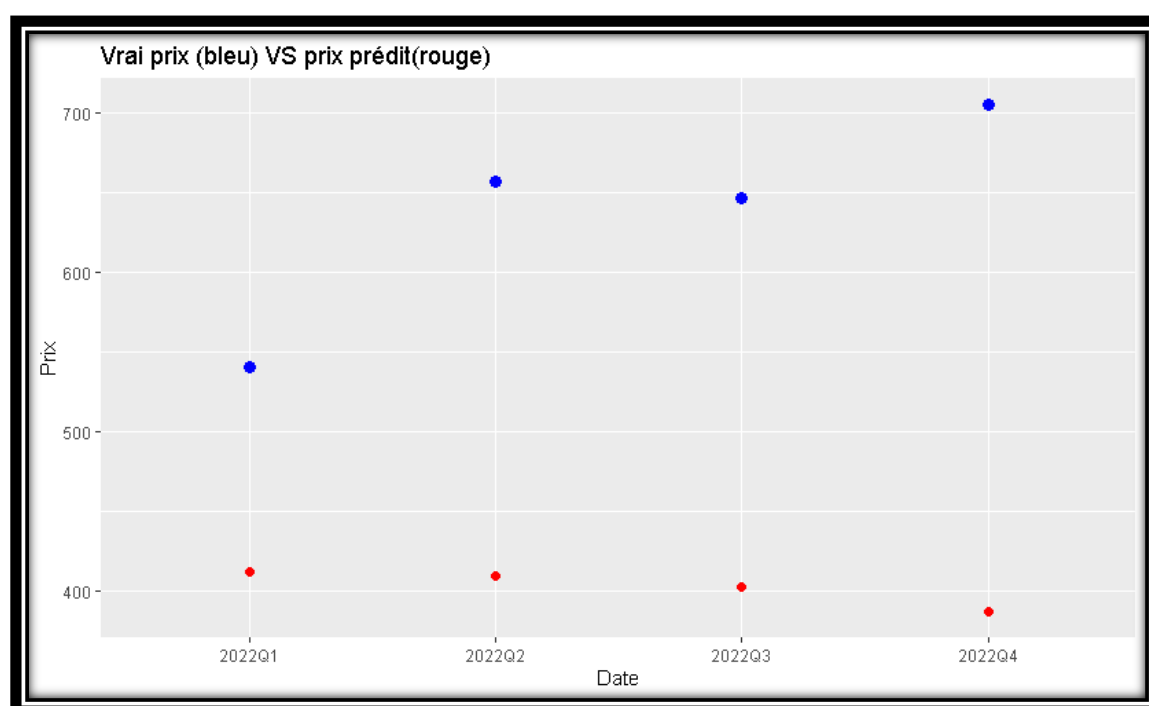
Pour appliquer le modèle complet, la même méthode sera suivie. Cependant, les variables explicatives comprendront cette fois l'entièreté des variables financières et ESG.

Lorsqu'on applique le random forest sur la base d'entraînement, on obtient maintenant un MSE de 814.72 et le pourcentage de la variance expliqué est de 88.56%. Ces résultats sont meilleurs, ce qui démontre que le modèle a mieux appris lorsque des variables ESG ont été rajouté.

Cependant, appliqué sur la base de test, le modèle affiche un RMSE de 244.62 ce qui reste relativement haut, mais surtout un R2 de -15.56. Le modèle n'est toujours pas meilleur qu'une simple prédiction de la moyenne des valeurs cible. Il y a donc un fort risque que le modèle ait sur-appris.

Résultat de la prédiction :

Dates	Vrai prix	Prix prédit
2022-T1	540.7	412.10
2022-T2	657.1	409.67
2022-T3	646.8	402.04
2022-T4	705.4	386.48

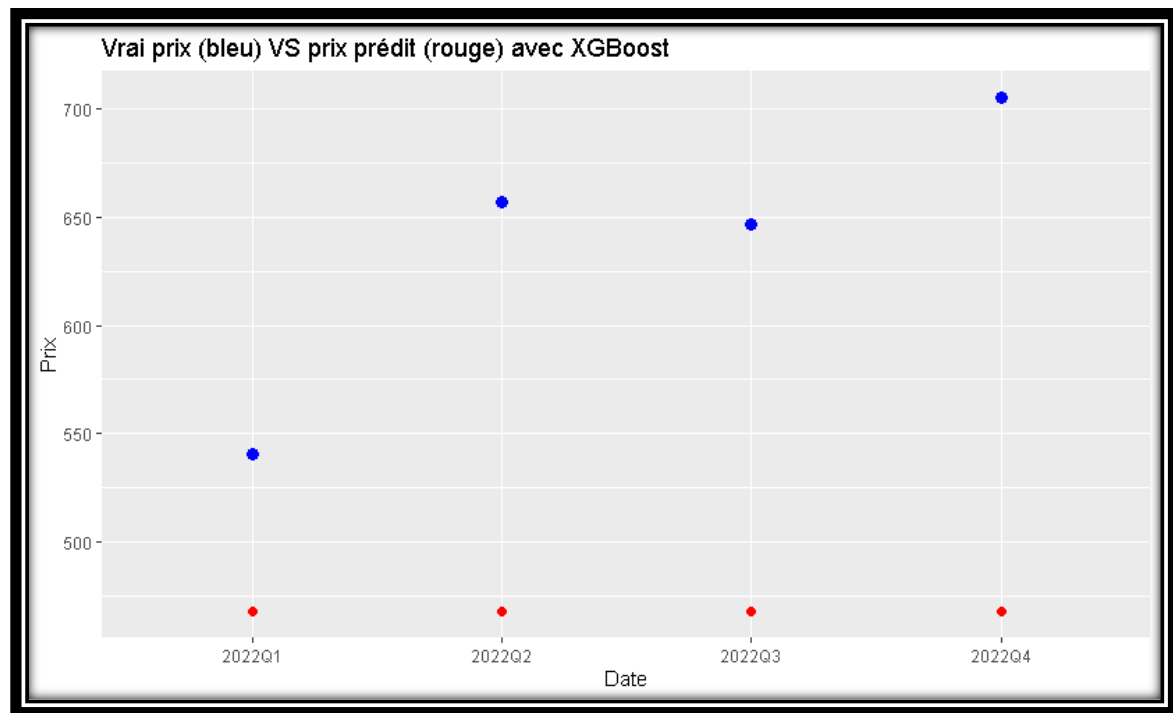


Les prix prédits restent bien loin des prix réels. On peut néanmoins noter une amélioration de la précision du modèle, notamment sur la base d'apprentissage.

Pour le XGboost, il faut encore appliquer directement le modèle entraîné sur la base de test pour voir sa précision. Il affiche un RMSE de 180.20 et un R2 de -7.9. Bien que toujours décevants, ces résultats montrent encore une amélioration du modèle lorsque l'on rajoute des variables ESG. Le R2 se rapproche de plus en plus d'un chiffre positif.

Résultat de la prédiction :

Dates	Vrai prix	Prix prédit
2022-T1	540.7	467.61
2022-T2	657.1	467.58
2022-T3	646.8	467.62
2022-T4	705.4	467.62



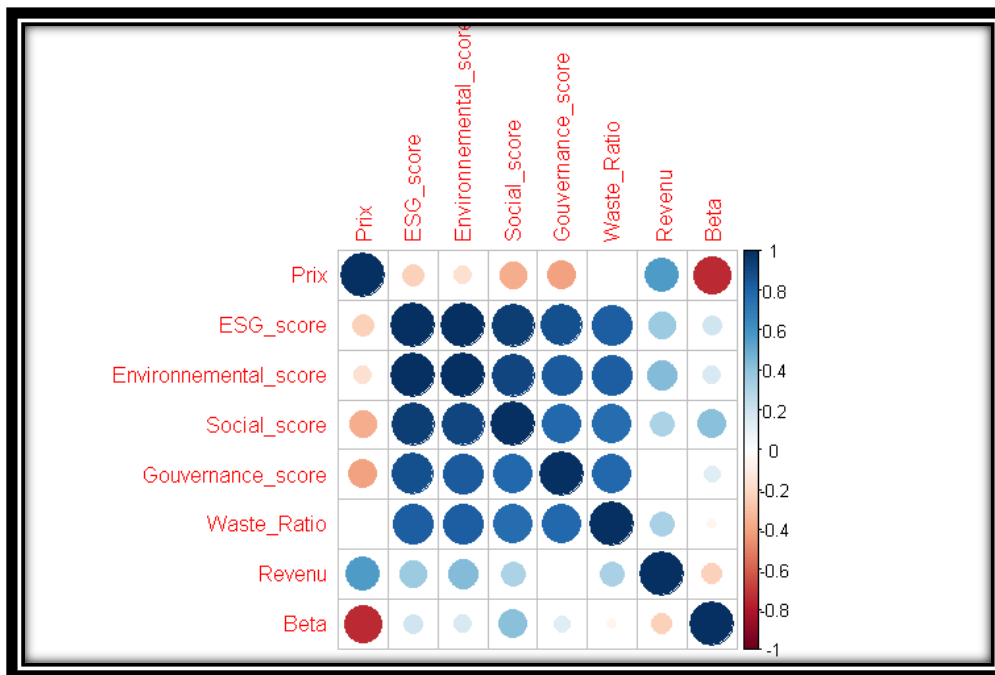
Le modèle arrive à être relativement précis pour le premier semestre, mais n'arrive pas à suivre la tendance haussière du prix de l'action.

Pour résumer le cas LVMH, les modèles arrivent difficilement à prédire correctement le prix futur de l'action. Il est fortement possible que les modèles sur-apprenne ou sous-apprenne sur la base de test. Cependant, les résultats semblent tout de même s'améliorer lorsque l'on rajoute des variables ESG quel que soit le modèle. De plus, la précision du random forest sur la base d'entraînement semble indiquer qu'en rajoutant ces données, le modèle arrive à capturer plus de lien entre les variables pour prédire le prix. Cela montre que, malgré les mauvais résultats, les données ESG sont (au moins légèrement) pertinente à utiliser dans un modèle quantitatif pour cette entreprise.

Cas de Total :

NB : Seul total sera rajouté pour compléter l'analyse pour éviter une redondance. En effet, le même code a été appliqué sur chaque entreprise, et les résultats des analyses ne diffèrent pas du cas d'une de ces deux entreprises.

La seconde entreprise à analyser est le cas de Total. Ses scores ESG sont bien supérieures à ceux de LVMH, et le prix de son action ne connaît pas une hausse aussi forte que LVMH.



Comme le montre le graphique, la matrice de corrélation est légèrement différente, puisque cette fois-ci, ce n'est pas la variable du ratio des déchets mais celle du Beta qui est corrélée négativement. Cela s'explique car Total a un Beta bien plus défensif que LVMH. De plus, ici encore les scores ESG sont corrélés entre eux et avec le prix, ce qui ne vient pas contredire notre analyse précédente.

Modèle financier :

Lorsque l'on utilise un random forest sur la base d'entraînement, celui-ci obtient un pourcentage de la variance expliqué de 55%. En revanche il est intéressant de notifier qu'après avoir effectué les tests de prédiction, on a un RSME de 6.41 et un R2 de -2.13. Donc bien que toujours décevant, les résultats sont tout de même supérieurs à ceux pour LVMH.

On retrouve cette amélioration avec le XGboost. Celui affiche un RMSE de 5.3 et un R2 de -1.14

Le modèle financier de Total semble meilleur que pour l'autre entreprise. Il est difficile de savoir exactement pourquoi, mais il est possible que le prix de l'action de Total soit bien plus affecté par les revenus de l'entreprise que pour LVMH. L'action de cette dernière pourrait dépendre d'autres facteurs qui ne sont pas pris en compte dans le modèle. Une autre possibilité est que la moyenne trimestrielle des prix de l'action de Total augmente plus faiblement que pour LVMH. Le modèle est donc naturellement plus proche.

Résultat des prédictions des 2 modèles :

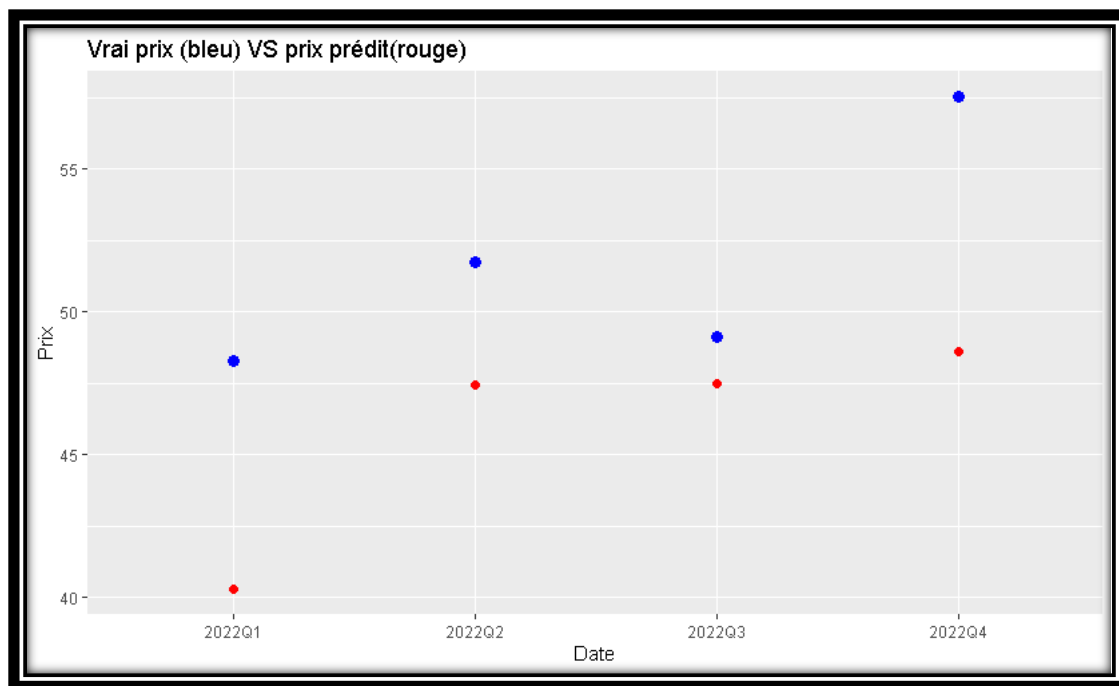
Random forest :

Dates	Vrai prix	Prix prédit
2022-T1	48.29	40.27
2022-T2	51.73	47.44
2022-T3	49.12	47.5
2022-T4	57.56	48.63

Xgboost :

Dates	Vrai prix	Prix prédit
2022-T1	48.29	43.15
2022-T2	51.73	48.38
2022-T3	49.12	51.12
2022-T4	57.56	49.12

Représentation graphique de la prédiction avec les deux modèles :

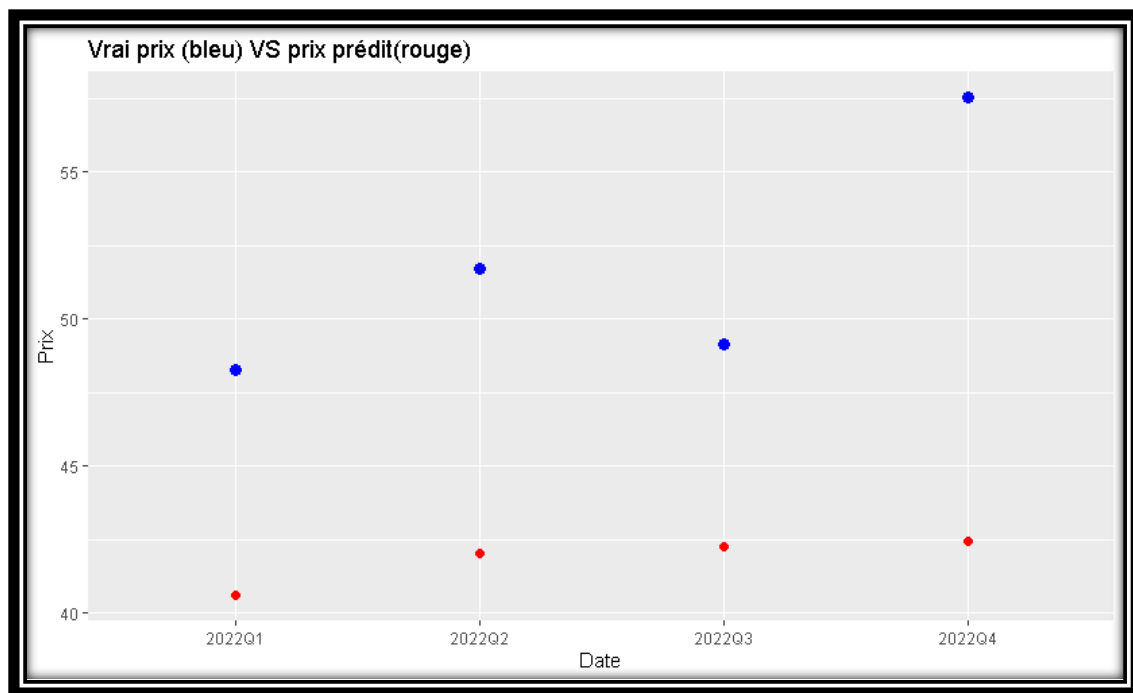


Modèle complet :

Avec l'ajout des variables ESG, le random forest affiche maintenant une variable expliquée de 66.64%. Il capte plus d'information qu'avec les variables purement financière. Cependant, à contrario de l'analyse précédente, le RMSE et le R2 diminue, passant 10.36 et -7.17

Résultat de la prédiction et graphique :

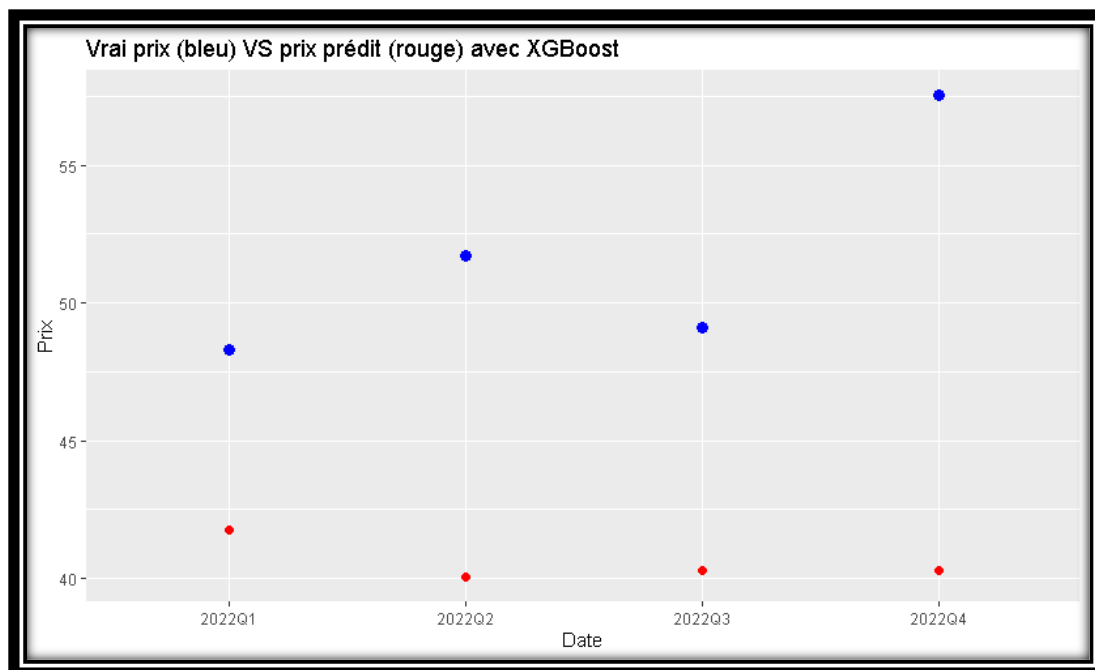
Dates	Vrai prix	Prix prédit
2022-T1	48.29	40.57
2022-T2	51.73	42.005
2022-T3	49.12	42.24
2022-T4	57.56	42.43



De façon identique, avec le XGboost il y a maintenant un RMSE de 11.77, et un R2 de -9.54.

Résultat de la prédiction et graphique :

Dates	Vrai prix	Prix prédit
2022-T1	48.29	41.77
2022-T2	51.73	40.02
2022-T3	49.12	40.30
2022-T4	57.56	40.30



L'analyse de Total vient donc contredire ce qui a été vu précédemment sur LVMH. Dans le cas de cette entreprise, l'ajout de variable ESG n'a pas permis une meilleure prédiction du prix de l'action. Pire encore, celle-ci est moins précise que pour le modèle financier simple.

Ce que montrent ces analyses :

Si l'ajout de variables ESG ont permis une meilleure prédiction avec LVMH, l'inverse c'est produit pour Total.

Là où la première entreprise connaissait une amélioration de ses notes ESG au fil des années, les scores pour la seconde stagnaient (voir régressé pour celui lié à la gouvernance). Il est donc possible que cette baisse soit la cause des mauvais résultats. Il est aussi possible que la cause soit dû aux prix de l'action de Total qui fluctuent, alors que ceux de LVMH sont croissant. Enfin, il est tout simplement possible que les facteurs ESG ne permettent pas une meilleure prédiction.

Conclusion :

Ainsi, après défini ce qu'est l'ESG et justifié le choix des variables sélectionnées pour cette analyse, les principales actions prises par les pouvoirs (qu'ils soient nationaux ou européens) ont été expliqués. L'application d'un véritable cadre légal, grâce à des normes et des labels, permet une transparence accrue et une véritable définition des actions étant des pratiques ESG. Les agents économiques, qu'ils soient des fonds de gestion ou des épargnants, ont une meilleure vision des actions des entreprises. Ce cadre permet ainsi la création d'un cercle vertueux pour des pratiques plus respectueuses de l'environnement et des populations. En effet, la transparence augmente la pression des populations sur les entreprises et incite les investisseurs à mieux choisir où est placé leurs argents. Les pratiques ESG réduisant plusieurs risques pouvant affecter négativement les entreprises, ces mêmes investisseurs ont alors une incitation d'autant plus grande à choisir les entreprises étant conformes à ces règles. De plus, les normes supranationales créent une standardisation des pratiques et mesures des actions ESG, ce qui permet une comparaison des différentes entreprises plus faciles à réaliser.

Ensuite, la partie quantitative a été abordée. L'étude des données, une étape cruciale dans la création de modèle, a montré que de nombreuses variables ont dû être modifiées pour pouvoir être utilisées. L'étude statistique a par la suite orienté le choix du modèle. Ce choix s'est porté sur un random forest et un XGboost, deux modèles relativement robustes au problème de corrélation entre les variables qui était présent. L'application de ces modèles sur deux entreprises, LVMH et Total, a démontré des résultats contraires. Les critères ESG ont certes permis une amélioration de la prédiction pour le prix des actions de LVMH, mais ils ont également baissé celle pour Total. Il est donc impossible de répondre positivement à la problématique, mais il est aussi impossible d'y répondre négativement. Il semblerait toutefois que l'ajout des variables ESG ont permis aux modèles de capter plus de lien entre les variables de la base d'entraînement qu'avec un simple modèle financier. Bien que ce résultat soit à prendre avec des pincettes, car pouvant simplement être dû à un surajustement du modèle, celui-ci pousse à l'optimisme pour réaliser une nouvelle analyse, avec des modèles différents ou mieux construits.

Limites et ouvertures :

Il existe néanmoins plusieurs limites à cette analyse. Tout d'abord, il se peut qu'une des hypothèses posées lors de l'introduction soit invalidée. Si les critères ESG actuels ne sont pas correctement mesurés, il est impossible de conduire une analyse pertinente. Ces données étant d'ailleurs mesurées de façon annuelle, il a fallu les transformer. Cette étape a pu modifier les liens, et donc fausser les modèles, mais interroge également sur le choix de la variable cible. Le prix de l'action étant calculé de façon journalière, l'utilisation d'une variable cible calculée également de façon annuelle (ou au moins trimestrielle) aurait réduit le nombre de transformation à effectuer.

De plus, les mesures étant plutôt récentes, la base de données n'est pas assez complète. Reconduire cette analyse dans dix ans, avec l'ajout des données de cette décennie permettrait de la rendre bien plus précise, permettant aux modèles un meilleur entraînement, et d'éviter un potentiel sur ou sous-ajustement.

Enfin, le choix d'un plus grand nombre de variables ESG pourrait améliorer l'analyse. Comme expliqué, le choix de ne pas supprimer de variables a été effectué. Cependant, l'ajout de plus de variables permettrait de seulement garder les plus importantes dans le modèle, ce qui le rendrait de facto plus précis.

Bibliographie :

- [Qu'est-ce qu'un critère extra-financier ? | AMF \(amf-france.org\)](#)
- [Bloomberg Europe](#)
- [Accord de Paris sur le climat — Wikipédia \(wikipedia.org\)](#)
- [Énergie non renouvelable — Wikipédia \(wikipedia.org\)](#)
- [Label ISR - Pour des placements durables et responsables \(lelabelisr.fr\)](#)
- [La nouvelle directive CSRD, tout savoir pour mieux s'y préparer | AMF \(amf-france.org\)](#)
- [La Taxonomie européenne : pièce centrale de la stratégie réglementaire européenne pour une finance durable \(carbone4.com\)](#)
- [Greenwashing — Wikipédia \(wikipedia.org\)](#)
- [Institute for Sustainable Investing | Morgan Stanley](#)
- [Risque climatique — Wikipédia \(wikipedia.org\)](#)
- [Shein : des parlementaires américains alertent sur la marque et le travail des Ouïghours - Le Parisien](#)
- [Rapport de l'AMF sur l'investissement socialement responsable \(ISR\) dans la gestion collective | AMF \(amf-france.org\)](#)
-

Table des matières

Sommaire :	2
Introduction :	3
1 – Les facteurs ESG : Définition, cadre légal et leurs impacts théorique sur la performance des entreprises.....	5
A) Définition et sélection des critères ESG pour l'analyse quantitative :	5
Choix des variables :	6
Choix des entreprises :	9
B) Cadre légal, normes et labels ESG : Synthèse et lien avec la performance financière.	11
Les greens bonds : leurs importances dans l'ESG et leurs liens avec la taxonomie européenne..	13
Le greenwashing : une limite importante de l'ESG :	14
Pourquoi les critères ESG permettent-ils une amélioration de la performance financière ?	14
Conclusion entre le développement d'un cadre ESG et la performance financière.	16
2 – L'analyse quantitative : Implémentation des variables et création du modèle.	18
A) Intégration des variables dans le modèle.....	18
B) Création du modèle et analyse des résultats.	23
Etude statistique :	23
Le non-choix de la régression linéaire multiple :	26
Le choix du modèle :	27
Cas de LVMH :	28
Cas de Total :	33
Ce que montre ces analyses :	37
Conclusion :	38
Limites et ouvertures :	39
Bibliographie :	40

