

iscte

INSTITUTO
UNIVERSITÁRIO
DE LISBOA

U LISBOA

UNIVERSIDADE
DE LISBOA

Análise econométrica aplicada à complexidade do Rating ESG

Joana Maria Carriço Marvanejo

Mestrado em Matemática Financeira

Orientadora:

Doutora Diana Elisabeta Aldea Mendes, Professora
Associada, ISCTE – Instituto Universitário de Lisboa

Setembro, 2023

Departamento de Finanças

Departamento de Matemática

**Análise econométrica aplicada à complexidade do Rating
ESG**

Joana Maria Carriço Marvanejo

Mestrado em Matemática Financeira

Orientadora:

Doutora Diana Elisabeta Aldea Mendes, Professora
Associada, ISCTE – Instituto Universitário de Lisboa

Setembro, 2023

Agradecimentos

O percurso do qual muito me orgulho e que percorri até hoje, fruto de muito trabalho, resiliência e persistência, que não teria sido o mesmo sem aqueles que dele fizeram parte e aos quais devo o meu sincero agradecimento.

A primeira palavra de agradecimento é dedicada, sem qualquer dúvida ou hesitação, aos meus pais e irmão. Obrigado pelo apoio incondicional que sempre me deram, não só no decorrer deste processo, mas, especialmente, em todos os momentos, bons ou maus, da minha vida. Para vós, um só obrigado é pouco.

Seguidamente, à minha orientadora, Professora Doutora Diana Elisabeta Aldea Mendes, agradeço a receptividade em relação aos ensinamentos de Econometria, bem como, a visão holística e epistemológica que conduziu este estudo.

Quero agradecer a toda a minha família, os votos de apoio e motivação que me acompanharam e tornaram a conclusão desta jornada possível.

Contudo, estou ainda grata ao meu companheiro de vida que conseguiu ter a amabilidade, gentileza, e cooperação comigo todos os dias desde o início ao fim deste projeto.

Por fim, mencionar também os meus amigos. Muitos foram os dias que, racionalmente, abduquei das suas companhias para acabar este projeto. Todavia, sem eles, nada disto faria sentido.

Resumo

Nos últimos anos, o interesse dos *stakeholders* registou um aumento significativo relativamente aos comportamentos socialmente responsáveis e no foco do *business sustainability* nas empresas, o que levou a que várias empresas incorporassem fatores sociais, ambientais e de *governance* (ESG) nas suas estratégias e decisões de investimento.

O renovado interesse conduziu ao recurso de agências de *rating ESG*, para uma avaliação da performance ESG, contudo esta não é padronizada dado que cada agência utiliza metodologias e estruturas diferentes. Dada a divergência de abordagens, esta investigação tem como propósito compreender como a escolha de uma agência particular pode impactar os resultados e conclusões de uma análise empírica, bem como a influência de cada fator no *rating ESG*.

Esta pesquisa utiliza dados provenientes das agências *Sustainalytics* e *Refinitiv*, em empresas cotadas publicamente na América do Norte, Europa, Ásia e Austrália em novembro de 2020 e abril de 2022.

Com o objetivo de explorar esse propósito realizou-se, essencialmente, uma análise descritiva, de regressão com fatores específicos (E, S, G, geográficos, financeiros e setoriais), de correlação e uma divisão de um conjunto de treino e teste, permitindo verificar a capacidade de generalização e a confiabilidade do modelo em contexto prático.

Esta investigação não proporciona apenas contribuições académicas significativas, mas também oferece informações de natureza prática para organizações que aspiram aprimorar as suas abordagens de sustentabilidade e aperfeiçoar as suas avaliações ESG.

Assim, as conclusões da análise da *Sustainalytics* revelam maior clareza em relação ao *rating ESG* fornecido pela *Refinitiv*.

Palavras-Chave: Finanças Sustentáveis, *Rating ESG*, Regressão Linear Múltipla, Agências de *Rating*, Análise de Dados, Metodologia Quantitativa.

Classificação JEL: G24, I31, M14

Abstract

In the last years, stakeholders' interest has witnessed a significant rise in socially responsible behaviors and a focus on business sustainability within companies. This has led several enterprises to incorporate social, environmental and governance (ESG) factors into their strategies and investment decisions.

The renewed interest has prompted the use of *ESG rating* agencies to assess ESG performance. However, standardization is lacking as each agency employs distinct methodologies and structures. Given the divergence in approaches, this investigation aims to comprehend how the selection of a specific agency can impact empirical analysis outcomes and conclusions, as well as the influence of each factor on *ESG ratings*.

This study utilizes ESG rating data from *Sustainalytics* and *Refinitiv* agencies for publicly listed companies across North America, Europe, Asia, and Australia in November 2020 and April 2022.

To explore this purpose, a primarily descriptive analysis, regression with specific factors (E, S, G, geographical, financial, and sectoral), correlation and a training-test data split were conducted. This allows for assessing the model's generalization capability and reliability in practical contexts.

This research contributes not only to academia but also offers practical insights for organizations seeking to enhance their sustainability approaches and refine their ESG assessments.

The findings from the *Sustainalytics* analysis notably illuminate greater clarity regarding the *ESG rating* provided by *Refinitiv*.

Keywords: Sustainable Finance, ESG Rating, Multiple Linear Regression, Rating Agencies, Data Analysis, Quantitative Methodology.

JEL Classification System: G24, I31, M14

Índice

Agradecimentos	iii
Resumo	v
Abstract	vii
1. Introdução	1
2. Revisão de literatura	5
2.1 <i>Rating</i>	5
2.2 <i>Definição da ascensão do investimento sustentável</i>	5
2.2.1 <i>Tendências e Legislação</i>	6
2.2.2 <i>Definição das siglas ESG</i>	6
2.3 <i>Análise do comportamento dos investidores em investimentos sustentáveis</i>	7
2.3.1 <i>Repartição de ativos nas regiões</i>	8
2.4 <i>Divergências no ESG</i>	9
2.5 <i>Relação entre o desempenho ESG e Financeiro</i>	10
2.6 <i>Importância das escalas de classificação ESG</i>	12
2.6.1 <i>Escalas de classificação</i>	12
2.7 <i>Agências de Rating</i>	12
2.7.1 <i>ESG na Refinitiv</i>	13
2.7.2 <i>ESG na Sustainalytics</i>	14
2.7.3 <i>Rating ESG</i>	14
3. Metodologia	16
3.1 <i>Regressão Linear Múltipla e o Método dos Mínimos Quadrados (OLS)</i>	17
4. Dados	23
4.1 <i>Descrição das variáveis</i>	23
4.2 <i>Tratamento dos dados</i>	25
5. Modelação e interpretação de resultados	29
5.1 <i>Modelo I: Fatores E, S, G e o Rating ESG</i>	29
5.2 <i>Variáveis dummy</i>	32

5.3 Modelo II: Fatores geográficos, financeiros, E, S, G e o Rating ESG.	33
5.4 Modelo III: Fatores financeiros, E, S, G e o Rating ESG	36
5.5 Modelo IV: Fatores E, S, G, financeiros, geográficos, setoriais e o Rating ESG ..	37
5.6 Testes estatísticos.....	39
5.7 Modelo V: Seleção das variáveis mais significativas	40
5.8 Multicolinearidade.....	41
5.9 Matriz de correlação	41
6. Avaliação e previsão.....	44
7. Considerações finais	47
Referências Bibliográficas	53
Anexos	59

Lista de Figuras

Figura 2.1- Categorias e pesos dos fatores E,S,G	14
Figura 4.1- <i>Box-plot</i> do <i>rating ESG</i> da <i>Sustainalytics</i> , em função da região	27
Figura 4.2- <i>Box-plot</i> do <i>rating ESG</i> da <i>Refinitiv</i> , em função da região	28
Figura 5.1- Matriz de Correlação.....	43
Figura 6.1- Validação cruzada <i>K-Fold</i> da <i>Sustainalytics</i>	44
Figura 6.2- Valores previstos vs valores reais da <i>Sustainalytics</i>	44
Figura 6.3- Validação cruzada <i>K-Fold</i> da <i>Refinitiv</i>	45
Figura 6.4- Valores previstos vs valores reais da <i>Refinitiv</i>	45

Lista de Tabelas

Tabela 4.1- Descrição das variáveis selecionadas.....	25
Tabela 4.2- Análise descritiva das principais variáveis	25

Lista de acrónimos

DSC- Desempenho Social Corporativo
ESG- Ambiental, Social e de <i>Governance</i>
ODS- Objetivos de Desenvolvimento Sustentável
OLS- Método dos Mínimos Quadrados
ONU- Organização das Nações Unidas
PIR- Princípios para o Investimento Responsável
RSE- Responsabilidade Social Empresarial
SRI- Investimento Socialmente Responsável

1. Introdução

A sustentabilidade é hoje um imperativo mundial, perante desafios complexos e em constante evolução a necessidade de abordagens aprofundadas e soluções inovadoras é mais evidente do que nunca. Nesse contexto, esta dissertação de mestrado tem como objetivo investigar e analisar o *rating ESG*, com a intenção de enriquecer o conhecimento e compreensão dentro deste domínio.

Os Princípios para o Investimento Responsável (PRI), com o apoio da ONU, têm vindo a ajudar a corrigir esta distorção, conduzindo uma maior clareza para a relevância financeira dos temas ESG¹, fornecendo um modelo para que a comunidade mundial de investimento contribua com o desenvolvimento de um sistema financeiro mais estável e sustentável.

As Instituições Bancárias desempenham um papel fundamental no desenvolvimento do Financiamento Sustentável, sendo este definido como o financiamento que, para além de considerar os aspetos financeiros tradicionais como a rentabilidade, a liquidez e a solidez financeira na avaliação dos projetos de investimentos das empresas, integra também aspetos não financeiros relacionados com os critérios ESG como o risco de transição, o risco físico, as condições de trabalho e as práticas de *governance*. Atendendo à importância deste tema, tornou-se urgente a definição e implementação de um novo índice, designado *rating ESG*, materializado através da integração dos critérios ESG nos processos de tomada de decisão.

O *rating ESG* é um indicador destinado à avaliação do desempenho de empresas em termos ambientais, sociais e de *governance*, capaz de fornecer informações relevantes aos investidores para que possam tomar decisões informadas e alinhadas com os seus princípios éticos e sustentáveis, a fim de reduzir riscos, ampliar o retorno financeiro e atender às expectativas dos beneficiários e clientes.

Este tema ainda que remonte aos anos 90, em pleno século XXI ainda gera bastante controvérsia. O seu principal objetivo é apoiar as empresas no processo de transição para uma economia mais verde e mais inclusiva, fornecendo *inputs* para melhorar o seu *rating ESG*, de modo a reorientar os fluxos de capital para estratégias de financiamento sustentável, assegurando a maximização de valor para os diversos *stakeholders*.

Todavia, apesar do crescente interesse na redução da emissão de poluentes para o meio ambiente, ainda não se conseguiu conter o seu avanço e os impactos negativos decorrentes da gestão imoral dos recursos, o que por sua vez tem causado repercussões significativas na economia global. Diante da crescente consciencialização sobre os problemas ambientais têm

¹ Acrónimo, em inglês, para *Environmental, Social and Governance* que corresponde às dimensões ambiental, social e do *governance* das organizações.

sido estabelecidas metas sustentáveis, sobretudo no âmbito empresarial, que representa a maior parcela dos impactos ambientais, sociais e de gestão em escala global.

Para avaliar o desempenho financeiro de uma empresa, tanto em termos absolutos quanto relativos, os analistas dispõem de uma série de dados financeiros reportados pelas empresas de forma recorrente e padronizada, seguindo normas contabilísticas de cada região. No entanto, quando se trata dos fatores ESG, não existem padrões universais para avaliar uma empresa em relação ao seu desempenho ESG e, muitas vezes a divulgação dessas informações não é obrigatória.

Segundo, *Fang, et al., (2022)* a *digitalização* emerge como uma ferramenta essencial para reduzir as assimetrias de informação nos processos transacionais entre gestores e colaboradores, investidores externos e gestores, promovendo a transparência e a responsabilidade. Além disso, os *social media* desempenham um papel significativo na minimização do risco de infrações de *governance* corporativa como a diluição de capital ou perda de direitos dos acionistas (*Dyck et al., 2008*), bem como infrações às regras de divulgação de informações (*Kouwenberg e Phunnarungsi, 2013*), contribuindo para o aprimoramento das pontuações G. De referir que segundo os autores, a digitalização influencia positivamente o ESG, facilitando as empresas a melhorar o *goodwill* e as pontuações S e G, no entanto, pode não influenciar o critério E.

Neste contexto, a relevância do *rating ESG* na área de matemática financeira está intrinsecamente ligada ao crescente interesse nos mercados financeiros, por investimentos sustentáveis e responsáveis, visando analisar as métricas e modelos estatísticos envolvidos na atribuição do *rating ESG*, bem como a sua relação com o desempenho financeiro das empresas.

Esta pesquisa visa investigar a fraca correlação entre os *ratings ESG das duas* agências, determinar um modelo de regressão que leve em conta cada fator do *rating ESG*, avaliar a incerteza associada à avaliação do *rating ESG* e a sua influência nos processos de investimento sustentável. Com base na compreensão desses aspectos espera-se fornecer *insights* valiosos para aprimorar a análise e utilização do *rating ESG*.

Dado que é um tema contemporâneo, surgiu a complexidade na organização da base de ideias iniciais desta dissertação, uma vez que decorre do facto de ser um campo com inúmeras teorias e investigações concluídas, em curso, inconclusivas ou conclusivas, o que gera o desafio de restringir informações suficientes, detalhadas, necessárias e verídicas. Posteriormente, ao reunir as informações consideradas relevantes para a análise e interpretação da questão científica desta dissertação, existiu a objeção em obter dados fiáveis, concretos e que abrangessem períodos temporais substanciais.

Posto isto, teve-se em conta uma análise e interpretação matemático-financeira do *rating ESG*, concluindo que a forma mais clara e cuidada para explicar as contradições do ESG era

realizar uma análise inicial de um artigo científico que tem como base fontes fidedignas e confiáveis. Deste modo, utilizou-se como referência a base de dados recolhida por *Erhart* (2022), realizando métodos e técnicas estatísticas diferentes das seguidas pelo artigo científico, de seguida construiu-se uma nova seleção e representação dos dados da amostra.

Por sua vez, a implementação da abordagem ESG por parte de empresas e organizações pode acarretar custos significativos. A presença de tais custos conduz a um ponto de partida crucial, conforme delineado por *Quivy e Luc Campenhout* (2008), que é essencial como etapa preliminar de todo o processo de investigação.

A questão de partida basilar para a estruturação de toda a investigação, foi a seguinte:

Até que ponto um rating ESG mais elevado influencia o desempenho das empresas?

Adicionalmente, emergem perguntas intimamente associadas à questão primordial, que serão abordadas ao longo desta investigação:

1) Como é que as variáveis de fatores E, S e G afetam o rating ESG das empresas?

Dado que o desempenho é definido pelas entidades responsáveis pelos ratings ESG, a investigação tem como objetivo explorar a relação entre esses fatores e compreender se o desempenho em áreas específicas, como sustentabilidade ambiental, responsabilidade social e práticas de *governance*, influenciam significativamente o rating ESG.

2) Qual o impacto das variáveis financeiras no *rating ESG*?

O estudo baseia-se em compreender se a situação financeira, a dimensão da empresa e a capacidade de distribuição de dividendos estão associados a *scores* mais elevados de ESG.

3) Há variações significativas nos *ratings ESG* entre diferentes áreas geográficas e setores de atividade?

Isso permitirá avaliar se os fatores regionais e setoriais impactam a abordagem de sustentabilidade e responsabilidade corporativa e se há variações nas prioridades ESG.

4) Qual é a relação entre o *rating ESG* e o desempenho financeiro das empresas?

A hipótese subjacente é que empresas com melhores *ratings ESG* também podem demonstrar um desempenho financeiro mais sólido.

Serão examinadas as correlações entre o *rating ESG* e variáveis financeiras, procurando entender se a responsabilidade corporativa está associada a melhores resultados financeiros.

5) Como é que os diferentes fatores têm impacto no rating ESG?

Através da construção de modelos de regressão será analisado o grau de explicação das variáveis independentes sobre as variações nas variáveis dependentes, o que

permitirá avaliar a importância relativa de cada variável na previsão do *rating ESG* e entender quais são as principais impulsionadoras do seu desempenho.

A dissertação pretende dar resposta a estas questões e está estruturada em sete capítulos, de acordo com uma lógica sequencial.

No primeiro capítulo encontra-se a Introdução, seguida a Revisão de Literatura, que fará a explicação sucinta dos conceitos de *rating ESG*, traçando a sua evolução histórica, juntamente com as legislações e regulamentações relacionadas, bem como a análise do comportamento dos investidores, divergências, relação entre o desempenho ESG e financeiro e a importância das escalas de classificação ESG, com foco na *Refinitiv* e *Sustainalytics*.

No capítulo seguinte, a Metodologia, explora a aplicação da regressão linear múltipla e do método dos mínimos quadrados (OLS) para a análise dos dados.

No capítulo quatro serão abordados os Dados, onde se concentra a descrição das variáveis estudadas e o tratamento dos dados utilizados na pesquisa.

O quinto capítulo é dedicado à Modelação e interpretação dos resultados, onde se apresenta vários modelos, incluindo diversos fatores com o *rating ESG*, testes estatísticos validação da robustez dos modelos, incluindo o estudo da multicolinearidade e a análise da matriz de correlação.

No capítulo subsequente, subdividiu-se o conjunto de dados em treino e teste, delineando o processo que permitirá a validação e previsão dos modelos construídos.

Por último, no capítulo sete, para além de um breve resumo das várias etapas da pesquisa, apresentam-se as principais conclusões, procurando estabelecer uma ligação crítica dos resultados apurados com a natureza dos dados em si. São ainda referidos os aspetos em que este trabalho pode contribuir para o desenvolvimento do conhecimento científico, apontando para as suas principais limitações, finalizando-se com a sugestão de algumas propostas de investigação futura.

É relevante notar que, embora a parte do código em *Python*, utilizado na condução deste estudo não tenha sido incluído integralmente nos anexos, estou à disposição para o disponibilizar completo mediante solicitação e fornecer o suporte necessário.

2. Revisão de literatura

2.1 Rating

O *rating* de uma empresa é uma avaliação ou classificação atribuída a ela por agências de *rating* ou instituições financeiras especializadas. Essa avaliação é baseada numa análise abrangente de aspectos financeiros, de crédito e de risco da empresa. O propósito do *rating* é proporcionar aos investidores, credores e outras partes interessadas uma visão da capacidade da empresa em cumprir as suas obrigações financeiras, bem como uma avaliação do seu perfil de risco.

Os *ratings* são normalmente representados por letras ou combinações de letras, como AAA, AA+, A-, B, entre outros (ver anexo A).

Cada agência de *rating* pode ter a sua própria escala de classificação e critérios de avaliação. Geralmente, as classificações mais altas indicam menor risco e maior capacidade de pagamento, enquanto as classificações mais baixas indicam maior risco e menor capacidade de pagamento.

Os *ratings* são amplamente utilizados pelos investidores e credores para tomar decisões informadas sobre investimentos e empréstimos. Eles podem influenciar as condições de empréstimos, taxas de juros, aceitação de títulos emitidos pela empresa e a confiança geral no desempenho financeiro da empresa.

No contexto do ESG, também existem agências de *rating* que avaliam o desempenho das empresas nessas áreas e atribuem classificações específicas.

2.2 Definição da ascensão do investimento sustentável

O investimento sustentável e os critérios ESG têm vindo a ganhar importância (ver anexo B) com a criação do Índice de Sustentabilidade *Dow Jones*. Este foi o primeiro índice global de ações a introduzir critérios ESG (Schmutz et al., 2020).

As empresas deste índice, não são apenas as maiores em termos de *market cap* mas também têm que cumprir alguns níveis de critérios para serem incluídas.

A publicação de *Ethical Investing, de Domini e Kinder (1986)* e *Investing for Good: Making Money While Being Socially Responsible de Lydenberg e Domini (1993)*, profetizou o que hoje chamamos de investimento sustentável. Conhecido na época como investimento ético ou investimento socialmente responsável, os autores defendiam uma abordagem que “examina os interesses e práticas comerciais relacionadas a preocupações éticas, como contratação de minorias, riscos ambientais e fabricação de armas, e explica como fazer investimentos rentáveis sem abdicar da ética pessoal”. O SRI, como ficou conhecido, visa capacitar indivíduos e consultores financeiros a alinhar decisões de investimento com valores pessoais

associados à visão de um mundo melhor, usado para avaliar empresas do ponto de vista da sustentabilidade (*Morningstar*, 2023).

Enquadrado no vocabulário das finanças tradicionais, o relatório foi um catalisador para converter a orientação de valores do SRI para um foco no *valor* - especificamente, no risco e oportunidades associados a fatores ESG no contexto dos deveres fiduciários de investidores institucionais (*Domini e Kinder*, 1986). E deu ao mundo uma nova sigla: *ESG*.

Esta vertente será cada vez mais escrutinada devido às crescentes pressões políticas e sociais e à existência de um maior número de dados monitorizáveis (*Informa D&B*, 2022).

2.2.1 Tendências e Legislação

Em 2004, o Pacto Global das Nações Unidas publicou um relatório endossado por 20 grandes instituições financeiras, juntamente com o Banco Mundial e a *International Finance Corp.* composto por recomendações sobre como integrar melhor as questões ESG.

Mais tarde, em 2006, a ONU apoiou os 6 PIR para incentivar as empresas a divulgar mais informações ESG (*Sievänen et al.*, 2013).

Em 2015, foi adotada uma abordagem mais aprofundada. Os governos dos 193 Estados membros da ONU assinaram a Agenda para o Desenvolvimento Sustentável 2030, que engloba os 17 ODS ([ver anexo C](#)), consistindo numa série de iniciativas, com três objetivos principais: erradicar a pobreza, proteger o planeta e corrigir desigualdades (*ONU*, 2015).

A ideia principal é que todos os intervenientes de mercado que contribuem para atingir os objetivos sustentáveis, possam vir a ser mais valorizados pelos investidores, que também terão mais certeza sobre o foco e o impacto dos seus investimentos (*Sætra*, 2021), surgindo assim nos últimos anos vários instrumentos legislativos com requisitos legais e de reporte de informação não-financeira.

Sachs (2015), argumenta que as pressões que encaminharam o ESG para o centro das atenções, relacionam-se, não só com a maior escassez de recursos e com os impactos dos desastres naturais, mas também com as crescentes expectativas de que as empresas cada vez mais se empenhem em atingir os ODS (*Elkington*, 2018), definidos pelas Nações Unidas.

2.2.2 Definição das siglas ESG

Uma avaliação ESG avalia o comprometimento de uma empresa quando se trata de questões ESG por outro lado, o *score ESG*, reflete a posição de uma entidade relativamente aos fatores ESG que a caracterizam num dado momento e quanto ao seu setor e a empresas de dimensão equiparável (*Informa D&B*, 2022). Dentro de cada uma dessas questões, há vários parâmetros a serem medidos, dependendo da empresa ou do setor.

Dentro dos *critérios E*, estes são indicadores ambientais que refletem a atuação de uma sociedade a lidar com questões de matéria ambiental, tanto no que respeita à política de preservação, quanto ao uso eficiente dos Recursos Naturais (gestão de produtos e serviços), como a emissões de gases de efeito estufa (GEE), poluição, resíduos e Incentivos Públicos para proteção ambiental (*Informa D&B, 2022*).

Conforme *Färber et al. (2020)*, o *fator S*, inclui indicadores sociais que mostram como uma sociedade assume as suas responsabilidades sociais e podem variar desde o Capital Humano (relações laborais, saúde e segurança, formação, direitos humanos e igualdade de género), o Envolvimento com Clientes (gestão de produto, de clientes e *data privacy*), o Envolvimento com Fornecedores (pagamento a fornecedores e avaliação do risco de *delinquency*²) e do Envolvimento com a Comunidade (donativos e incentivos públicos de inclusão social).

Por fim, dentro dos *critérios G*, que são indicadores relacionados com a *governance* das entidades, nomeadamente o *Corporate Governance* (gestão, transparência empresarial, diversidade na gestão, conformidade empresarial e interesse dos *stakeholders*) e a Sustentabilidade do Negócio (indicadores de desempenho, sustentabilidade e resiliência financeira³, certificações e distinções) (*Informa D&B, 2022*).

Constata-se que o investimento sustentável tem atraído um especial interesse, entre investidores e instituições, criando assim uma nova realidade financeira empresarial, destacando a necessidade de uma análise aprofundada das consequências ambientais, sociais e de *governance* (*Bak et al., 2017*).

2.3 Análise do comportamento dos investidores em investimentos sustentáveis

Hoje, o investimento sustentável abrange uma série de estratégias, tendo como premissa a visão de que: todo o investimento tem efeitos no mundo real, sejam eles positivos ou negativos.

A sustentabilidade começa a ser um tema sobre o qual as entidades financeiras recolhem a opinião dos investidores. Por isso, diversos dados, pesquisas e estudos são publicados com frequência e atualizados a cada momento, contribuindo para uma maior evolução desta megatendência.

² Risco de *delinquency* (Modelo preditivo) - Indicador que reflete a probabilidade de nos próximos 12 meses, uma entidade registar um atraso de pagamentos superior a 90 dias face aos prazos acordados.

³ Resiliência financeira - Indicador que traduz a resiliência financeira de cada empresa relativamente ao seu setor, medida através da ponderação da estrutura de custos, de endividamento da empresa, da sua margem de lucro e rendibilidade, e da eficiência que regista no seu investimento, medindo assim a capacidade de uma empresa enfrentar e reagir a choques excecionais no seu processo produtivo e/ou comercial.

De acordo com um artigo feito pela KPMG (*Swinburne et al, 2019*), o interesse e foco dos investidores, não se limita apenas aos dados financeiros, mas atualmente, também se centra em aspetos não financeiros, tornando-os cada vez mais importantes. O que distingue os investidores sustentáveis, dos investidores focados exclusivamente em resultados financeiros é a sua intenção de integrar a análise de riscos e oportunidades ESG, abordar os efeitos negativos das atividades de mercado e obter impactos positivos nos seus negócios (*Informa D&B, 2022*). Portanto, a maioria dos investidores espera que a integração do ESG leve a maiores recompensas a longo prazo (*Maiti, 2020*).

Embora *Friede et al. (2015)* tenha descoberto que quase 90% dos estudos ilustram uma relação não negativa entre o desempenho do ESG e o *corporate financial*, a integração do fator ainda reúne dúvidas em alguns investidores.

No estudo *Schroders Global Investor Study 2022* ([ver anexo D](#)), verificou-se que "dois terços dos portugueses inquiridos (66%) acreditam que o investimento sustentável é a única forma de assegurar a rentabilidade a longo prazo, em comparação com 60% das pessoas a nível global" (*FundsPeople, 2022*).

Mas esta não é a única conclusão sobre o tema. Segundo o estudo, os portugueses inquiridos estão cada vez mais conscientes de que as práticas de investimento têm um papel importante a desempenhar na abordagem de questões de sustentabilidade global. Cerca de 67% da amostra concorda que "o investimento pode impulsionar o progresso em desafios de sustentabilidade, tais como as alterações climáticas". Adicionam também que as alterações climáticas, precisamente, são motivo de preocupação pelo impacto que podem causar nos seus investimentos. Neste estudo, 58% da amostra classifica esta preocupação como "fundamental" para eles.

Deste modo, como conta a *Schroders*, 70% dos investidores especialistas partilham a opinião de que "investir de forma sustentável pode gerar mudanças positivas quando se trata de desafios como as alterações climáticas" (*FundsPeople, 2022*).

2.3.1 Repartição de ativos nas regiões

No final de 2019, o total de ativos líquidos (TNA) registados foi de 48,8 mM€ em todo o mundo, constatando um crescimento significativo em comparação a 2008, quando o TNA registou 13,65 mM€ (*Inverco, 2018*).

Os dados da *Morningstar (2022)* mostram que, dos quase 2,8 tr USD investidos globalmente em fundos sustentáveis, quase 2,3 tr estavam na Europa. Isso representa 82% do volume total de ativos. Os Estados Unidos representam 12%, enquanto os restantes 6% estão divididos entre a Ásia-Pacífico (3%), a Oceânia (1%), o Japão (1%) e o Canadá (1%).

Na verdade, a Europa não é apenas líder em investimento socialmente responsável em termos de ativos sob gestão, mas o continente também lidera em termos de fluxos e número de fundos, segundo a *Global Sustainable Investment Alliance* (2020).

Em termos de número de produtos, a Europa é a região com o maior número de fundos ESG registrados. No final de março de 2022 havia quase 5000 fundos ESG, representando 77% do total. Os Estados Unidos tinham 555 (9%), enquanto na Ásia-Pacífico, Oceânia, Japão e Canadá o número de estratégias comercializadas continuava a ser muito escasso.

Os fluxos de fundos sustentáveis, notavelmente o *ETF iShares Paris-Aligned Climate MSCI USA*,⁴ foi lançado apenas em fevereiro de 2022, mas garantiu o quarto lugar entre os líderes de fluxos sustentáveis, atraindo 608 M USD, durante o primeiro trimestre de 2022 (*Morningstar*, 2022).

Nos últimos dois anos, desde a pandemia, os mercados cresceram tremendamente, contudo, com o crescente aumento da volatilidade e da inflação, os investidores começaram a diversificar os seus portfólios (*Morningstar*, 2022).

Olhando para o futuro, a regulamentação continuará a ser um tema importante para os investidores voltados para a sustentabilidade em 2023, começando o ano em perceber como as estratégias de investimento sustentável são percebidas e avaliadas pelos investidores.

2.4 Divergências no ESG

As classificações ESG das diferentes agências divergem, introduzindo incertezas nas decisões e criando um desafio, para os tomadores de decisão (*Berg, Kölbel e Rigobon*, 2022).

Existem diversas pesquisas relacionadas com a avaliação e classificação do ESG e todas apontam para a *falta de convergência* no método de medição, conduzindo à necessidade de maior consistência e transparência entre as agências, obtendo uma maior qualidade no método de avaliação (*ERM*, 2019).

Como resultado, esses estudos argumentam que os investigadores reconhecem a necessidade de apoiar a comunidade de finanças sustentáveis a alcançar os objetivos autoimpostos por meio da medição ESG e contribuir para a uniformidade do método de classificação, como apontado por *Dorflleitner, Halbritter e Nguyen* (2015) e *Drempetic, Klein e Zwergel* (2020).

Berg, Kölbel e Rigobon (2022), apontam para a relevância do problema de divergência na avaliação do *rating ESG*, referindo que as agências de *rating* medem o mesmo, usando diferentes indicadores, pesos e conjunto de atributos. Os autores concluíram que a divergência dos indicadores utilizados na medição, explica mais de 50% da divergência geral.

⁴ <https://www.ishares.com/us/products/325725/ishares-paris-aligned-climate-msci-usa-etf/>

Yu e Luu (2021), utilizaram a divulgação dos scores ESG de dados da *Bloomberg* e perceberam que as características das empresas explicam a maior parte da variação do ESG, enquanto as variações nos fatores do país, como corrupção e direitos políticos, têm menos influência.

Dada a natureza subjetiva dos critérios, verifica-se uma variedade de definições de ESG, sendo esta a principal razão para levar as agências de *rating* a atribuir diferentes scores para a mesma empresa (Sorrosal-Forradellas et al., 2023).

Em primeiro lugar, as agências podem discordar sobre como medir o mesmo fator pois, apesar dos esforços de diversas organizações de padronização, não há uma abordagem universalmente aceita para medir indicadores não financeiros (Grewal, 2019). As agências de *rating* utilizam centenas de variáveis relacionadas ao ESG provenientes de relatórios da empresa e arquivos regulatórios e, portanto, devem ser consistentes entre as mesmas. No entanto, muitas outras são obtidas por meio de entrevistas ou questionários e relatórios independentes de terceiros com abordagens potencialmente discrepantes.

Ao analisar os relatórios de Responsabilidade Social Empresarial (RSE), Kiri e Nozaki (2020), utilizaram o *text mining* para avaliar as atividades ESG das empresas. Por outro lado, em vez de analisar relatórios RSE, ou seja, informações fornecidas pela própria empresa, Azhar et al (2019), realizaram uma análise de conteúdo de artigos de notícias, como fonte de dados que mitiga problemas como a sub ou super-representação de tópicos ESG/ RSE. Hisano et al (2020), também utilizou notícias como fonte de informação, argumentando, assim como Azhar et al (2019), que elas são menos suscetíveis à manipulação, desenvolvendo uma abordagem baseada em redes com o objetivo de prever a inclusão de uma empresa numa lista de exclusões de investimentos, devido a questões ESG.

Assim, o objetivo atual do estudo consiste em apostar na crescente inovação do tema, expandindo a discussão da divergência da classificação ESG dos diferentes índices da bolsa de valores e tratar *benchmarks* tradicionais da mesma forma que os *benchmarks* ESG, contribuindo para aumentar a transparência em todo o universo de índices, conforme recomendado pela Comissão Europeia em 2019.

Assim, a metodologia utilizada por diferentes agências para decidir quais os indicadores relacionados com o ESG a considerar e como agregá-los numa pontuação final, será sempre alvo de discussão.

2.5 Relação entre o desempenho ESG e Financeiro

Estudos empíricos que investigam a relação entre a pontuação ESG e o desempenho financeiro não apresentam um consenso. Vários estudos concluem que existe uma relação positiva entre as práticas ESG e o desempenho financeiro (Waddock e Graves, 1997; Fischer e Sawczyn, 2013; Friede et al., 2015; Velte, 2017; Alareeni e Hamdan, 2020). Em contraste,

outros estudos não encontram relação significativa, negativa ou não-linear (Nollet et al., 2016; Xie et al., 2019; Duque-Grisales e Aguilera-Caracuel, 2021).

Waddock e Graves (1997), refere que o DSC está *positivamente* associado ao desempenho financeiro anterior e futuro, concluindo que os dados suportam teorias que afirmam uma *relação positiva* entre os recursos excedentes e o DSC e uma boa gestão e o DSC assim, os autores consideram que os resultados obtidos são uma espécie de "círculo virtuoso". Além disso, Fischer e Sawczyn (2013), também encontraram evidências de que a relação é afetada pelo nível de inovação da empresa.

Segundo Velte (2017), com base em dados ESG retirados do banco de dados Thomson Reuters Datastream, constata que a *relação positiva* no desempenho financeiro se mantém tanto para a pontuação ESG total, quanto para os seus três componentes, à semelhança do estudo de Alareeni e Hamdan (2020), que apurou que a divulgação de ESG por uma empresa tem um *efeito positivo* sobre o desempenho operacional, financeiro e de mercado de empresas americanas. Além disso, ao contrário de Velte (2017), o artigo *não encontrou evidências de uma relação unidirecional* entre os três pilares e o desempenho financeiro.

Pelo contrário, Alareeni e Hamdan (2020), concluíram que os pilares ambiental e social afetam *positivamente* o desempenho financeiro, enquanto o pilar de *governance* tem um *efeito negativo*. Além disso, a conclusão geral do artigo é que quanto maior o nível de divulgação de ESG, maior o desempenho financeiro.

Por outro lado, Duque-Grisales e Aguilera-Caracuel (2021), verificam uma *relação linear negativa* no desempenho de ESG sobre o desempenho financeiro. Além disso, a conclusão geral de Duque-Grisales e Aguilera-Caracuel (2021), apoia que uma pontuação elevada de ESG leva a um desempenho económico pior, tanto para o ESG total quanto para os três pilares individuais.

Apesar dos resultados contraditórios, Nollet et al. (2016), observou uma *ausência de relação linear significativa* entre a pontuação de ESG e o desempenho financeiro. No entanto, o estudo encontra uma *relação quadrática*, o que implica que a RSE só compensa após um certo montante de investimento e objetivos, ou seja, antes que o ponto crítico seja atingido, os investimentos em RSE terão um impacto negativo sobre o desempenho financeiro.

À semelhança anterior, Xie et al. (2019), salientou uma *relação não linear* entre a eficiência corporativa e as estratégias de RSE concluindo que, para empresas com pontuações moderadas de ESG, as práticas de ESG estão positivamente associadas ao desempenho financeiro. No entanto, para empresas com pontuação ESG baixa ou alta, o desempenho ESG prejudica a eficiência corporativa.

Aferindo-se que existem várias teorias em relação ao desempenho financeiro.

2.6 Importância das escalas de classificação ESG

O investimento na sustentabilidade e nos fundos sustentáveis está em potencial crescimento, de acordo com as classificações ESG os fundos supramencionados, têm atingindo entradas consideráveis (Hartzmark e Sussman, 2019). As classificações ESG têm vindo a tornar-se parte integrante de decisões financeiras, empresariais e de consumo, levando os investidores a valerem-se das mesmas para obter uma avaliação de terceiros sobre o desempenho ESG das empresas.

A sua importância tem sido reconhecida por novos líderes políticos, na Europa, a presidente da Comissão Europeia, *Ursula Von Der Leyen*, que apresentou o *European Green Deal* em 2019, comprometeu a Europa a atingir a neutralidade carbónica em 2050, seguidamente também nos EUA, o presidente eleito, *Joe Biden*, se comprometeu com 2 bi USD, em gastos com a sustentabilidade em 2020 (Erhart, 2022).

Assim, as agências de *rating ESG* tornaram-se instituições influentes nos mercados financeiros atuais, afetando uma ampla gama de decisões importantes.

2.6.1 Escalas de classificação

Segundo *Orecchio* (2022), as escalas de classificação ESG são uma forma de avaliar a performance das empresas em relação a fatores ESG. Essas escalas são usadas por investidores e gestores de ativos para identificar empresas que apresentam boas práticas em ESG, o que pode indicar um bom desempenho financeiro e uma dedução de riscos a longo prazo.

As escalas ESG são geralmente baseadas em critérios e indicadores definidos por organizações especializadas como a *Global Reporting Initiative (GRI)*, *Sustainability Accounting Standards Board (SASB)* e a *Principles for Responsible Investment (PRI)*.

Esses critérios podem incluir, desde a gestão de emissões de gases de efeito estufa, até diversos ODS. Estas escalas são uma ferramenta importante para os investidores que pretendem integrar fatores ESG nas suas estratégias de investimento, podendo usar essas escalas para comparar o desempenho ESG de diferentes empresas e setores, identificar riscos e oportunidades associados às questões ESG e tomar decisões de investimento mais informadas e conscientes.

2.7 Agências de Rating

As agências de rating ESG que avaliam e analisam as práticas de negócios relacionadas com o meio ambiente, social e *corporate governance* das empresas (García et al.,2019), surgiram para atender à elevada procura de informações por investidores que procuram investir em empresas sustentáveis.

Exemplos dos sistemas de scores ESG mais importantes são o *Sustainalytics ESG*⁵, o *Bloomberg ESG Disclosure Score*⁶, o *Thomson Reuters ESG Score*⁷ e o *MSCI ESG Score*⁸, entre outros.

Embora as agências de classificação ESG não divulguem ao público as suas metodologias completas para avaliar empresas, nesta dissertação, é fundamental conhecer mais sobre as metodologias adotadas pela *Sustainalytics* e *Refinitiv*, cujos dados serão utilizados nesta pesquisa.

2.7.1 ESG na Refinitiv

Esta dissertação usa dados ESG da *Refinitiv* (anteriormente *Thomson Reuters ASSET4*), cobrindo cerca de 85% do valor de mercado global, em mais de 630 medidas ESG diferentes, com histórico que remonta a 2002 (*Refinitiv, 2022*). As classificações estão disponíveis para mais de 12000 empresas públicas e privadas em todo o mundo sendo que, na maioria dos casos, os dados ESG são atualizados uma vez por ano, de acordo com a divulgação ESG das próprias empresas (*Refinitiv, 2022*).

A metodologia por detrás da classificação ESG da *Refinitiv* é completamente diferente da utilizada pela *Sustainalytics*. Em vez de medir o risco, a classificação ESG da *Refinitiv* é projetada para avaliar o desempenho, compromisso e eficácia ESG de uma empresa em 10 categorias principais, como: direitos humanos, comunidade, inovação, recursos ambientais, emissões ambientais, responsabilidade pelo produto, acionistas, gestão e estratégia de RSE (*Refinitiv, 2022*).

Por sua vez, reformula as pontuações dos três pilares e a pontuação final ESG, onde é um reflexo do desempenho, comprometimento e eficácia ESG da empresa. Segundo a *Refinitiv (2022: 3)*, "as pontuações finais são baseadas no desempenho relativo dos fatores ESG, do setor da empresa (para ambiental e social) e do país de incorporação (para *governance*)". Adicionalmente, a *Refinitiv* classifica os *ratings ESG* numa escala de 0-100, em que 0 (100) é a pontuação ESG mais baixa, ou seja, a pior (mais alta, logo a melhor).

Reitera-se que a pontuação total ESG representa a soma relativa dos pesos das categorias E, S e G. Por sua vez, a [Figura 2.1](#), é a prova visual de como isso pode ser representado, demonstrando o problema de divergência dos pesos das diferentes categorias.

Esses dados são direcionados para medir, de forma objetiva e transparente, o desempenho, o comprometimento e a efetividade relativa de uma empresa em relação ao ESG, com base em dados reportados pela empresa ([ver anexo E](#)).

⁵ <https://www.sustainalytics.com/esg-data>

⁶ <https://www.bloomberg.com/professional/product/esg-data/>

⁷ <https://www.refinitiv.com/pt/sustainable-finance/esg-scores>

⁸ <https://www.msci.com/our-solutions/esg-investing/esg-ratings>



Figura 2.1- Categorias e pesos dos fatores E,S,G

Fonte: Refinitiv

2.7.2 ESG na *Sustainalytics*

A *Sustainalytics*, propriedade da *Morningstar*, é uma fornecedora global de análises e classificações ESG que tem como objetivo ajudar os investidores a integrar fatores de sustentabilidade nas suas decisões de investimento e avaliar como os riscos ESG podem afetar o desempenho do investimento de longo prazo (Garz e Volk, 2018). Assim, esta classificação pretende medir o potencial impacto das questões ESG no valor da empresa.

Com um universo de mais de 16000 empresas, incluindo públicas e privadas, e em mais de 40 setores, a *Sustainalytics* utiliza o seu extenso banco de dados, para analisar e medir os riscos E, S, G, aos quais uma empresa está exposta, bem como comparar o desempenho de gestão com seus *peers* e *scores* do setor (Garz e Volk, 2018). Assim, obtém uma melhor compreensão das melhores práticas do setor e de ferramentas inovadoras para desenredar novas opções de sustentabilidade, enquanto identifica oportunidades de mitigação de riscos (ver anexo F).

À semelhança da agência anterior, a *Sustainalytics*, também classifica os *ratings ESG* numa escala de 0-100, exceto em que 0 (100) é a pontuação ESG mais baixa, ou seja, a melhor (mais alta, logo a pior), existindo a necessidade de transformar a escala de classificação na mesma da *Refinitiv*.

2.7.3 Rating ESG

Como observado por Singh et al. (2012), é crucial ter um método adequado que possa levar a um resultado fidedigno.

Na prática, cada agência de *rating* aplica os seus próprios critérios de desempenho social, e a sua metodologia de avaliação, o que pode levar a uma certa controvérsia devido à incerteza, ambiguidade ou imprecisão (Liern e Pérez-Gladish, 2018).

Reiterando o *rating ESG* possibilita:

“i) Avaliar e hierarquizar as empresas em termos de sustentabilidade, segundo uma determinada escala de *rating*, podendo contribuir para a definição de *pricing* do crédito;

ii) Conjugar esta avaliação de aspetos não financeiros, com a avaliação financeira tradicional (dividida entre *ratings investment grade* e *ratings non-investment grade*);

iii) A transparência na solução, pois permite identificar as componentes de sustentabilidade que uma empresa pode melhorar;

iv) A elevada granularidade na vertente de empresas analisadas e na vertente de dados ponderados, pois facilita a seleção de parceiros de negócio com base na sua sustentabilidade;

v) A objetividade na avaliação das empresas, pois simplifica a análise da sustentabilidade das carteiras de clientes e fornecedores;

vi) O complemento de resposta a requisitos de regulação” (CGD,2022).

3. Metodologia

A presente dissertação tem como objetivo essencial investigar a influência de diversos fatores no *rating ESG* das empresas. Para alcançar esse objetivo utilizou-se, como base, os dados inicialmente recolhidos do artigo de *Erhart (2022)* com referência a duas agências, a *Refinitiv* e a *Sustainalytics*.

Os dados foram concentrados de forma sistemática, considerando uma amostra diversificada de empresas de diferentes setores e localizações geográficas, adotando uma abordagem quantitativa, permitindo uma análise estatística. Consequentemente, a amostra do estudo foi selecionada com base na disponibilidade dos dados e na representatividade das empresas de diversos setores e regiões, abrangendo cerca de 60% das ações nos índices analisados, uma vez que atualmente não há classificação ESG disponível para as restantes 40% das empresas na amostra dos 20 índices de ações.

O foco principal desta dissertação é alcançar uma série de objetivos interligados. Primeiramente, pretende-se investigar e avaliar a existência de uma possível correlação fraca entre os *ratings ESG* atribuídos por diferentes agências, destacando as variações nas avaliações. Adicionalmente, procura-se a construção e validação de um modelo de regressão robusto, considerando minuciosamente cada componente, a fim de estabelecer um modelo explicativo sólido para o desempenho ESG e financeiro das empresas.

De seguida, uma das metas cruciais é a avaliação das diferenças associadas à atribuição dos *ratings ESG* e a sua repercussão nas estratégias de investimento sustentável. Este estudo explora como a incerteza inerente à avaliação do *rating ESG* pode influenciar as decisões dos investidores que procuram integrar considerações ambientais, sociais e de *governance* nos seus portfólios. Por fim, o estudo visa examinar de que maneira a escolha da agência pode afetar a incerteza percebida e, concludentemente, como as hipóteses e regressões podem revelar a preponderância dessa escolha nas análises.

Por sua vez, este estudo adotou também uma abordagem exploratória, utilizando dados secundários recolhidos de artigos científicos e relatórios de agências de *rating*, permitindo analisar as relações entre os diferentes fatores e o *rating* de maneira mais ampla e abrangente.

Foram utilizadas técnicas de regressão para avaliar a relação entre os diversos fatores e o *rating ESG*, considerando diferentes variáveis independentes. Foram assim desenvolvidos os seguintes modelos para os dados de cada uma das duas agências consideradas:

- Modelo 1: Fatores E, S, G e o *Rating ESG*;
- Modelo 2: Fatores geográficos, financeiros, E, S, G e o *Rating ESG*;
- Modelo 3: Apenas fatores financeiros, E, S, G e o *Rating ESG*;
- Modelo 4: Fatores E, S, G, financeiros, geográficos, setoriais e o *Rating ESG*;

- Modelo 5: Variáveis mais significativas com base na significância individual;

De referir que, durante todo o processo de pesquisa, foram observadas considerações éticas para garantir a confidencialidade e a privacidade das informações das empresas analisadas. Além disso, todas as fontes foram citadas e referenciadas devidamente para preservar a integridade acadêmica da pesquisa.

Embora existam algumas limitações nesta pesquisa, como a disponibilidade dos dados, a utilização de apenas duas agências, a divergência e a particularidade do tema, a análise desses dados pode fornecer informações valiosas sobre a relação entre as classificações ESG e o desempenho financeiro, contribuindo para uma melhor compreensão das implicações do *rating ESG* no contexto dos investimentos financeiros e socialmente responsáveis. Essas limitações podem afetar a generalização dos resultados, porém foram consideradas no processo de análise e interpretação.

Esta pesquisa contribuirá para uma melhor compreensão das implicações do *rating ESG* e abre caminho para futuras investigações nesse campo promissor.

3.1 Regressão Linear Múltipla e o Método dos Mínimos Quadrados (OLS)

Neste Capítulo são abordados os principais aspetos deste modelo, nomeadamente, a expressão analítica e os pressupostos dos resíduos. A facilidade de uso e a simplicidade conceitual dos Modelos de Regressão Linear são alguns dos motivos que justificam a ampla adoção e a diversidade de aplicações desses modelos em contextos quotidianos.

O objetivo deles é estudar a existência de uma relação linear entre uma ou mais variáveis independentes e uma variável alvo.

Neste âmbito, considere-se $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})$ um vetor de p variáveis aleatórias (independentes) e denomine-se por Y_i a variável dependente.

O modelo de regressão linear múltipla define-se por:

$$Y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i \quad (1)$$

em que β_0 representa o valor médio da variável dependente quando $x_i = (0, 0, \dots, 0)$ e $\beta_1, \beta_2, \dots, \beta_p$ representam a variação média da variável resposta a cada incremento unitário na respetiva variável independente (*ceteris paribus*) e ε_i representa a variação aleatória da variável em torno de um plano de regressão. Todos os coeficientes de regressão, β , assumem valores reais, a estimar pelo método dos mínimos quadrados.

É possível concluir também que, dado um vetor $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})$ de observações das variáveis independentes, se tem que:

$$E(Y_i) = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} \quad (2)$$

$$VAR(Y_i) = \sigma^2 \quad (3)$$

O primeiro passo na aplicação deste modelo consiste em obter as melhores estimativas para os chamados coeficientes de regressão, $\beta_0, \beta_1, \beta_2, \dots, \beta_p$. Como tal, considere-se:

$$X = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1p} \\ 1 & x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}, Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \text{ e } E = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (4)$$

três matrizes, em que cada linha da matriz X corresponde a uma de n observações de um vetor x_i , cada célula y_i corresponde ao valor observado da variável dependente respetiva e a mesma linha do vetor E corresponde ao erro aleatório associado a essa observação. Nesta notação, pode-se escrever o conjunto das observações como uma equação matricial, isto é:

$$Y = X\beta + E, \quad (5)$$

em que:

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_p \end{bmatrix} \quad (6)$$

A partir do leque de possíveis variáveis explicativas, foram testadas várias combinações para a estimação dos parâmetros do modelo, sendo que a estimação foi realizada pelo método OLS.

O método dos mínimos quadrados baseia-se numa metodologia que procura encontrar o melhor resultado para um conjunto de dados, tentando minimizar a soma dos quadrados das diferenças entre os valores estimados e os valores observados. A esta diferença dá-se o nome de resíduos.

No fundo, este método resulta de um estimador que minimiza a soma dos quadrados dos resíduos da regressão obtida, por forma a maximizar o ajustamento do modelo com os dados observados.

$$\min \sum_{i=1}^n e_i^2 = \min \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (7)$$

Em que y é o valor observado e \hat{y} é o valor predito.

Para regressões lineares múltiplas, com mais de uma variável independente, a equação OLS é expandida para incluir todos os coeficientes correspondentes a cada variável independente:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_n x_n \quad (8)$$

Nesta equação, \hat{Y} é a variável dependente prevista, $\hat{\beta}_0$ é o coeficiente de intercetação, $\hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_n$ são os coeficientes das variáveis independentes. O objetivo do método OLS é

encontrar os valores de $\hat{\beta}_0, \hat{\beta}_1, \hat{\beta}_2, \dots, \hat{\beta}_n$, que minimizam a soma dos quadrados dos resíduos, ou seja, que melhor ajustam o plano aos dados observados.

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (9)$$

No decorrer da metodologia, a linguagem de programação *Python* terá um papel fundamental para conduzir a análise proposta. A biblioteca *pandas* será utilizada para organizar, limpar e preparar os dados, assegurando uma manipulação eficiente das informações, enquanto que a biblioteca *scikit-learn* será empregue para construir e validar os modelos de regressão, permitindo uma exploração detalhada das relações entre os fatores ESG e o *rating* das empresas.

É importante destacar que todas essas etapas são conduzidas com o auxílio da linguagem de programação, permitindo a realização de cálculos estatísticos complexos, a visualização de dados, testes de pressupostos e validação rigorosa dos modelos de regressão linear múltipla.

Tal como se segue, o processo de análise será dividido nas seguintes etapas:

- I. **Análise das Estatísticas Descritivas:** Realizar uma análise detalhada das estatísticas descritivas dos *ratings* ESG das duas agências, permitindo a compreensão de características fundamentais dos dados, identificação de tendências iniciais e a observação de uma base sólida, para as etapas subsequentes da análise.
- II. **Construção e Validação de Modelos de Regressão:** Desenvolver e validar modelos de regressão que estabelecem relações entre os fatores ESG, desempenho financeiro, geográfico e setorial, com o objetivo de criar modelos robustos capazes de capturar as interações complexas entre essas variáveis, fornecendo uma compreensão mais profunda das influências ESG. Seguida, de uma abordagem sistemática e rigorosa que envolve diversas etapas críticas, a seguir apresentadas:

(a) Análise Gráfica Inicial: Conduzir uma análise visual das relações entre as variáveis independentes e a variável dependente. Por meio de gráficos de dispersão, vai ser possível identificar padrões, tendências iniciais e possíveis *outliers* nos dados. Essa análise preliminar oferece *insights* valiosos sobre a natureza das relações entre as variáveis, auxiliando na compreensão inicial do fenômeno em estudo.

(b) Pré-processamento dos Dados: Realizar etapas fundamentais, como a limpeza dos dados para remover informações indesejadas ou incorretas, a

imputação de valores omissos, análise de *outliers* que possam distorcer as análises e transformação das variáveis categóricas em variáveis *dummy*, garantindo que todos os dados estejam prontos para análises estatísticas robustas.

(c) Estimação dos Parâmetros: Utilizar o método OLS, para estimar os parâmetros do modelo de regressão. Cada coeficiente de regressão representa a mudança média na variável dependente associada a uma unidade de mudança na variável independente, mantendo todas as outras variáveis constantes. Esses coeficientes são fundamentais para entender as relações quantitativas entre as variáveis e para formular previsões.

(d) Pressupostos dos Resíduos: Verificar os pressupostos dos resíduos para garantir a validade do modelo:

- **Normalidade dos Resíduos ($\varepsilon_i \sim N(0, \sigma^2)$):** Verificar se os resíduos se aproximam de uma distribuição normal por meio de gráficos de probabilidade normal (*QQ-plots*) e testes estatísticos, como o teste de normalidade de *Jarque-Bera* e *Shapiro-Wilk*. A normalidade dos resíduos é crucial, pois permite inferências estatísticas válidas e interpretação apropriada dos coeficientes estimados.
- **Homoscedasticidade ($VAR(\varepsilon_i) = \sigma^2$):** Verificar com testes estatísticos de *Breusch-Pagan* e de *White* e visualmente, por meio de gráficos de dispersão dos resíduos em relação aos valores ajustados ou às variáveis independentes, a validação que a variância dos resíduos seja constante em diferentes níveis das variáveis independentes.
- **Independência dos Resíduos ($E(\varepsilon_i, \varepsilon_j) = 0$):** Examinar os resíduos para detetar qualquer autocorrelação serial, ou seja, correlação entre resíduos de períodos adjacentes. O teste de *Durbin-Watson* é comumente usado para verificar a autocorrelação dos resíduos, em que valores próximos a 2 indicam independência dos resíduos, evitando que as conclusões sejam distorcidas.
- **Ausência de Relação entre Resíduos e Variáveis Independentes ($E(\varepsilon_i, X_j) = 0$):** Validar esse pressuposto por meio de gráficos, em que se as concentrações dos pontos forem próximas a zero e a linha se apresentar mais ou menos horizontal, sugere uma ausência de uma relação sistemática entre os resíduos e as variáveis

independentes, ou seja, que as variáveis independentes não estão a influenciar de forma previsível os resíduos.

- **Média Zero dos Resíduos ($E(\varepsilon_i) = 0$):** Este pressuposto implica que, em média, os resíduos (ou seja, as diferenças entre os valores observados e os valores previstos pelo modelo) devem agrupar-se em torno de zero, indicando que o modelo não tende a superestimar ou subestimar consistentemente os valores observados.

(e) Validação do Modelo: A validade do modelo é avaliada por meio de métricas estatísticas, sendo o coeficiente de determinação (R^2) uma das principais, que permitirá avaliar o grau de ajuste dos modelos aos dados, em que SQR é a Soma dos Quadrados de Regressão e SQT é a Soma dos Quadrados Total.

$$R^2 = \frac{SQR}{SQT} \quad (10)$$

III) Testes Estatísticos t e ANOVA: Realizar testes estatísticos t e análise de variância (ANOVA) para avaliar a significância estatística das variáveis independentes em relação ao *rating ESG*. Essa análise estatística fornecerá *insights* sobre quais fatores ESG possuem maior influência estatística no *rating*.

$$F = \frac{\frac{R^2}{k}}{\frac{1-R^2}{n-(k+1)}} = \frac{MSS(\text{Variância entre grupos})}{MSE(\text{Variância dentro dos grupos})} \quad (11)$$

Onde n representa o número total de observações em todos os grupos e k o número de grupos ou categorias que estão a ser comparados no teste de ANOVA.

- IV) Seleção das Variáveis mais relevantes:** Identificar as variáveis mais relevantes para os modelos de regressão específicos das duas agências, determinando quais os fatores ESG possuem um impacto estatisticamente significativo no *rating* de cada agência.
- V) Avaliação da Multicolinearidade:** Analisar a presença de multicolinearidade entre as variáveis independentes nos modelos de regressão, calculando os fatores de inflação de variância (VIF), valores elevados de VIF indicam correlações fortes entre as variáveis, o que pode afetar a interpretação dos coeficientes. Essa avaliação é essencial para assegurar a validade das análises estatísticas e a interpretação correta dos coeficientes estimados.

$$VIF_k = \frac{1}{1-R_k^2} \quad (12)$$

VI) Matriz de Correlação: Realizar uma análise da matriz de correlação entre as variáveis para identificar relações entre as variáveis independentes e o *rating*

ESG. Essa etapa é crucial para identificar e compreender como as diferentes variáveis estão correlacionadas e podem influenciar conjuntamente na construção dos modelos de regressão e do *rating ESG*.

VII) Conjunto Treino-Teste: Dividir o conjunto de dados num conjunto de treino e de teste, permitindo avaliar a eficácia dos modelos de regressão desenvolvidos, contribuindo para verificar como os modelos se comportam em dados não utilizados durante o treino, garantindo a generalização dos resultados.

VIII) Previsão e Avaliação da Performance: Prever valores futuros da variável dependente com base nas variáveis independentes, compondo um modelo final. A análise da correlação entre os valores previstos e reais ajuda a avaliar a precisão e a qualidade das previsões geradas pelo modelo.

A metodologia estruturada aqui descrita, assegura a robustez e confiabilidade das análises realizadas no estudo das relações entre os diferentes fatores e *ratings ESG*.

4. Dados

A partir deste momento, apresentam-se análises gráficas e tabelas que foram geradas por meio das ferramentas de análise de dados no *Python*. Estes recursos visuais representam resultados diretos do processo de análise e desempenharão um papel fundamental na geração de *insights* ao longo deste relatório.

4.1 Descrição das variáveis

O conjunto de dados inicial consiste numa amostra de 33 variáveis (numéricas e categóricas) de natureza *cross-section data* (ver anexo G).

Os dados contêm apenas os meses de novembro de 2020 e abril de 2022, de forma a incorporar um período impactado pela pandemia Covid-19, evidenciando a necessidade de priorizar critérios sociais, saúde e segurança levando as empresas a reavaliarem as suas práticas e políticas de responsabilidade corporativa e outro período marcado pelo conflito da Guerra Rússia-Ucrânia, repercutindo preocupações geopolíticas na avaliação de *governance* e nos fatores de risco associados a determinadas empresas, afetando a economia mundial e os diferentes setores.

O conjunto da agência da *Sustainalytics* contém 1990 observações (1016 em novembro de 2020 e 974 em abril de 2022) por outro lado a *Refinitiv* apresenta uma amostra com 1692 observações (722 em novembro de 2020 e 970 em abril de 2022).

Como referido anteriormente na secção 2.7.3 as agências de *rating ESG*, possuem uma forma de classificação distinta, pelo que se realizou uma normalização de escalas, convertendo a escala da *Sustainalytics* para a da *Refinitiv*, ou seja 0 é o pior e 100 o melhor.

A análise econométrica terá como objetivo verificar se e como o *rating ESG* é afetado pelas restantes variáveis escolhidas. Em resumo, o propósito do estudo é compreender se com um conjunto reduzido de dados também se podem retirar conclusões corretas e suficientes acerca dos dados ESG.

Por conseguinte, selecionaram-se as variáveis dependentes e independentes, ou seja, do universo das 33 variáveis, escolheram-se 12 variáveis independentes e 2 variáveis dependentes. Na [Tabela 4.1](#), são descritas as variáveis consideradas ao longo desta dissertação.

Sendo assim, a variável dependente a considerar é o *rating ESG* (“Sus_ESG”, “R_ESG”), de cada uma das agências, em percentagem, sendo uma forma de avaliação que mede o desempenho de empresas dos 20 índices da bolsa de valores, em relação aos critérios ambientais, sociais e de *governance*. Demonstra ser um dos fatores mais importantes que mede as atividades sustentáveis de um conjunto de empresas, surgindo como um indicador representativo do desenvolvimento sustentável.

As variáveis independentes são: o fator E (“Sus_E”, “R_E”), em percentagem, que representa a pontuação de sustentabilidade ambiental, ou seja, que mede a contribuição de uma empresa para a proteção e preservação do meio ambiente. O fator S (“Sus_S”, “R_S”), em percentagem, que indica o nível de risco associado aos aspetos sociais de uma empresa. O fator G (“Sus_G”, “R_G”), em percentagem, mede a qualidade e eficácia das práticas de *governance* de uma empresa, incluindo transparência, responsabilidade corporativa e ética nos negócios.

A variável “*Index*”, categórica, representa o índice da bolsa de valores, uma medida estatística que representa o desempenho agregado de um grupo de empresas cotadas numa determinada bolsa de valores. Esses índices são calculados com base no valor de mercado das empresas, no número de ações negociadas e outros fatores relevantes, usados para monitorizar e avaliar o desempenho geral do mercado de ações. Na presente amostra existem 15 índices de ações europeus (Áustria, Bélgica, Dinamarca, Finlândia, França, Alemanha, Hungria, Itália, Países Baixos, Noruega, Rússia, Espanha, Suécia, Suíça, Reino Unido), 2 da América do Norte (Estados Unidos, Canadá), Austrália e 3 da Ásia (Hong Kong, China e Japão).

A variável “*BookValue*”, numérica, representa o valor contabilístico de uma empresa, podendo ser usada como uma métrica para avaliar a posição financeira coletiva das empresas que compõem o índice. No entanto, é importante observar que o “*BookValue*” não tem em consideração o valor de mercado das empresas nem fatores como as perspetivas futuras.

A variável “*MarkCap*”, numérica, representa o valor de mercado de uma empresa que é calculado através da multiplicação do preço atual das ações da empresa pelo número total de ações em circulação. As empresas com um “*MarkCap*” maior são geralmente consideradas de maior porte e podem ser vistas como mais estáveis e maduras, sendo um indicador importante, pois reflete a avaliação de mercado da empresa pelos investidores. Além disso, o “*MarkCap*” não tem em consideração fatores como dívida e outros ativos não negociados publicamente, o que pode afetar a avaliação total de uma empresa.

A variável “*DivYield*”, numérica, é uma medida financeira que expressa a relação entre o dividendo anual distribuído por uma empresa e o preço atual das suas ações. Portanto, quanto maior o “*DivYield*”, maior o retorno em termos de dividendos que o investidor pode esperar receber em relação ao preço da ação.

A variável “*Region*”, categórica, representa a área geográfica que, neste caso é apenas na Europa, América do Norte, Ásia e Austrália.

Por fim, a variável “*PeerGroup*” representa o setor de atividade de cada empresa.

Parte-se da ideia de que das variáveis escolhidas, “Sus_E”, “Sus_S”, “Sus_G”, “R_E”, “R_S”, “R_G”, vão ter uma forte relação de causalidade com o “Sus_ESG” e “R_ESG”, respetivamente.

Tabela 4.1- Descrição das variáveis selecionadas

Nome das variáveis	Tipo	Variável	Data extração
Sus_ESG (%)	Numérica	Dependente	Março 2023
Sus_E, Sus_S, Sus_G (%)	Numérica	Independente	
R_ESG (%)	Numérica	Dependente	
R_E, R_S, R_G (%)	Numérica	Independente	
PeerGroup	Categórica	Independente	
Index	Categórica	Independente	
Region	Categórica	Independente	
MarkCap	Numérica	Independente	
DivYield	Numérica	Independente	
BookValue	Numérica	Independente	

Fonte: Autor

4.2 Tratamento dos dados

Primeiramente, realizou-se uma análise descritiva dos dados, (Tabela 4.2) utilizando-se medidas estatísticas simples como o mínimo, o máximo, os quartis, a média, a mediana, a moda e o desvio padrão, com o intuito de sintetizar os dados numéricos. Essas medidas fornecem uma visão geral das características centrais, variabilidade e distribuição dos dados.

Executou-se uma análise das estatísticas descritivas das variáveis, sendo útil para identificar os valores atípicos, detetar erros ou inconsistências nos dados, obtendo uma compreensão inicial da distribuição dos dados e explorando tendências ou relações preliminares entre as variáveis.

Tabela 4.2- Análise descritiva das principais variáveis

	Sus_ESG	Sus_E	Sus_S	Sus_G	R_ESG	R_E	R_S	R_G
count	1990.000000	1990.000000	1990.000000	1990.000000	1692.000000	1692.000000	1692.000000	1692.000000
mean	72.804977	81.394703	69.314883	59.455440	66.281596	66.156638	67.706693	68.932606
std	12.092712	17.733557	13.290092	14.521585	17.654546	21.713379	19.872386	18.318651
min	0.000000	0.000000	0.000000	0.000000	3.537647	0.000000	0.000000	1.000000
25%	65.476190	71.119163	60.591133	51.664876	55.047181	54.000000	56.217664	57.249951
50%	73.886329	87.117552	70.443350	62.406015	69.000000	71.311572	72.000000	72.172465
75%	81.566820	95.096618	78.390805	69.656284	80.000000	82.173070	82.524091	83.000000
max	100.000000	100.000000	100.000000	100.000000	100.000000	98.000000	98.000000	98.000000

Além disso, utilizaram-se técnicas gráficas como histogramas (ver anexo H), gráficos de dispersão e diagramas de caixa, para visualizar e identificar correlações/padrões visuais nos dados, por exemplo observar se existe alguma relação significativa, negativa, não-linear ou nenhuma relação.

Através da tabela acima, verifica-se que, em média, o *rating ESG* e os fatores E, S e G da agência de *rating Sustainalytics*, são superiores à média da *Refinitiv*, à exceção da componente de *governance*.

Em média, na *Sustainalytics*, as empresas apresentam riscos G (59.5) mais baixos e fatores E (81.4) mais altos, o que pode indicar que as entidades gerem melhor os riscos ambientais do que os de *governance*. Isto deve-se ao facto da crescente regulamentação e legislação para a redução de comportamentos poluentes, o que pode influenciar positivamente o *rating* final. Adicionalmente o *rating ESG* apresenta um valor mínimo e máximo entre 0 e 100, respetivamente, o que traduz que há empresas com uma classificação ESG muito baixa ou muito alta, sendo a sua média de 72.8, o que comprova que a *Sustainalytics*, em média, apresenta melhores classificações ESG do que a *Refinitiv* (66.3).

Em contraste, na *Refinitiv*, as empresas têm, em média, um melhor desempenho em questões G (68.9) e pior nos fatores E (66.2). Esta discrepância pode surgir pela forma como as empresas estão a relatar informações sobre esses pilares, nomeadamente em termos de transparência empresarial, interesses dos acionistas, e como as agências as classificam segundo uma metodologia específica, diferenciada e mais focada na transparência e divulgação, concluindo-se que as empresas podem ser melhor avaliadas. Reforçando este facto, existe uma complexidade na questão G, pois existem diferenças culturais e regulatórias, podendo variar em diferentes países e setores, ou seja, o fator G pode não receber a mesma ênfase ou ser regida por padrões menos/mais rigorosos, o que pode levar a uma gestão deficiente/eficiente nessa área.

Com base na [Tabela 4.2](#), o 1º quartil, indica que o valor da variável “Sus_E” é aproximadamente 71.1 o que significa que apenas 25% das empresas possuem uma pontuação inferior a 71.1, no entanto a variável “R_E” assume um valor inferior, de cerca de 54.0, o que reforça que as classificações da componente ambiental são melhores na *Sustainalytics*. Já o 2º quartil, representa a mediana, em que a variável “Sus_ESG” é aproximadamente 73.9, o que indica que metade das empresas possui uma pontuação de sustentabilidade superior a 73.9, por sua vez a mediana na variável “R_ESG” é de 69.0, fortalecendo a mesma evidência.

Adicionalmente, com base na mesma fonte, observou-se uma distribuição dos dados com possíveis divergências, evidenciadas pelos *valores omissos*, deste modo optou-se por remover algumas variáveis do modelo inicial, criando assim um próprio modelo, que visa explicar a relação de cada fator E, S, G no *rating ESG* de cada agência incorporando apenas dois períodos de tempo.

Posto isto, foi levado a cabo um procedimento de remoção dos valores omissos das variáveis que constituem mais de metade das observações totais. Para tratar os valores omissos, primeiramente, optou-se por efetivar o método da imputação múltipla uma vez que

é uma técnica avançada que envolve a criação de várias estimativas para os valores omissos, com base em modelos estatísticos, combinando essas estimativas num resultado final. A ideia por trás da imputação múltipla é que os valores omissos não são estimados apenas uma vez, mas sim várias vezes para capturar a incerteza associada a essas estimativas. Contudo, o método contém desvantagens, como o pressuposto que exista uma certa distorção dos resultados da análise, da redução da precisão das conclusões, bem como a razão da manifestação dos valores omissos, que podem ser devido a um erro de medição, a uma não classificação ESG ou perda de dados.

Deste modo, conclui-se que a melhor abordagem seria pelo método da redução da amostra uma vez que a exclusão das observações não introduz um enviesamento significativo nos resultados ESG, obtendo-se uma redução na amostra para 1579 observações.

Através dos gráficos *box-plot* (Figura 4.1 e 4.2), visualiza-se, de forma mais clara, a distribuição estatística de cada uma das variáveis dependentes por meio de quartis, bem como uma maior capacidade de observação de *outliers*, com base nos valores mínimos.

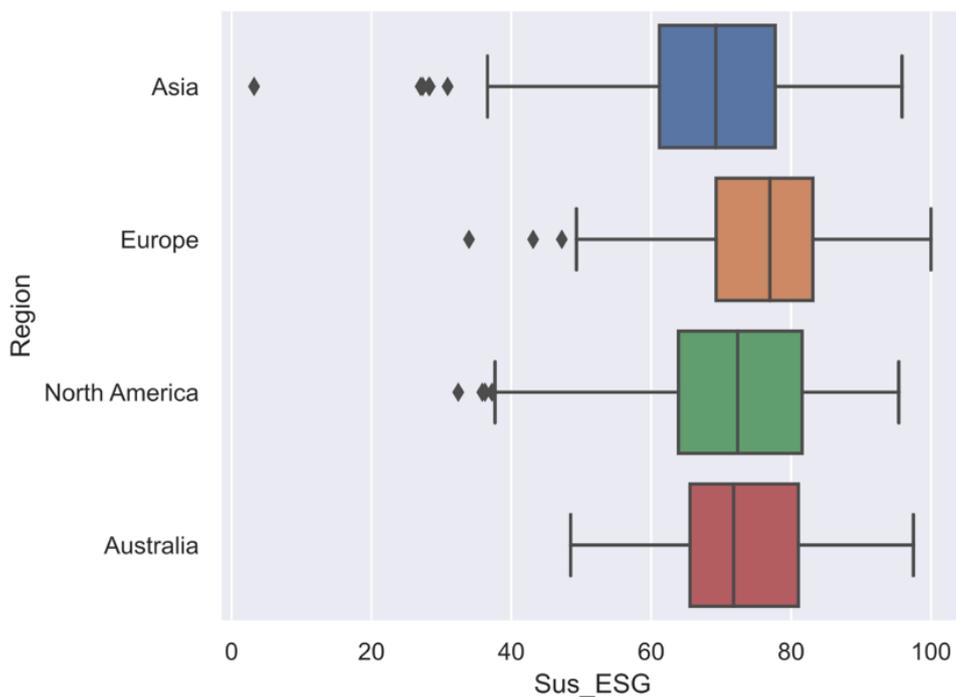


Figura 4.1 - Box-plot do rating ESG da Sustainalytics, em função da região

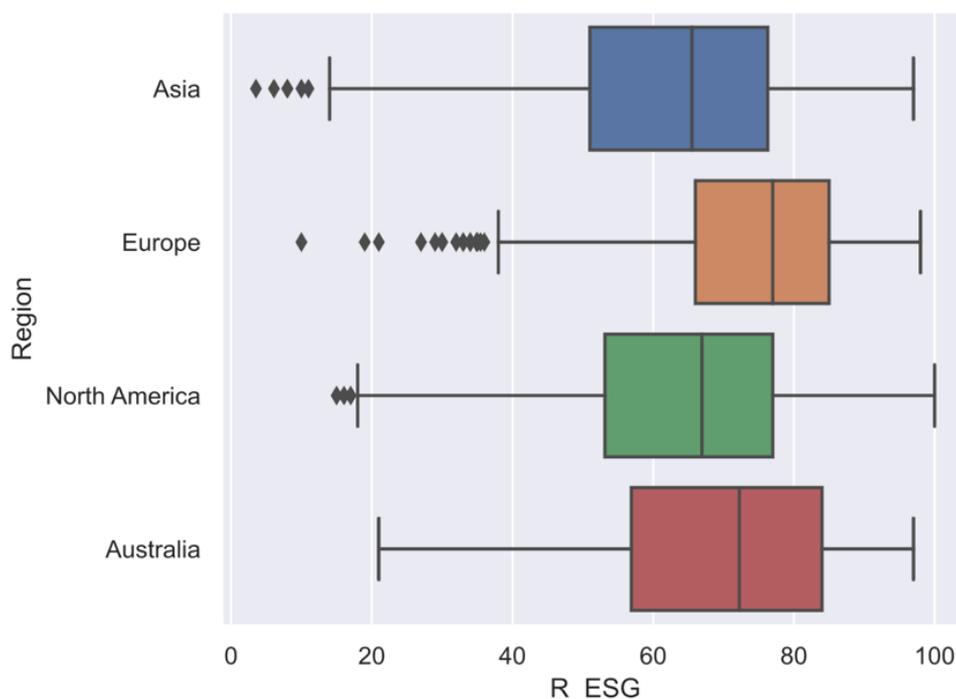


Figura 4.2 - Box-plot do rating ESG da Refinitiv, em função da região

Optou-se por não remover os *outliers* da amostra visto que poderia enviesar os resultados positivamente uma vez que os mesmos possuem uma classificação menor no *rating ESG*, por sua vez ainda se presume que os *ratings ESG* da *Sustainalytics* (Figura 4.1) sejam melhor classificados do que a *Refinitiv* (Figura 4.2), pois esta possui uma maior distribuição de classificações abaixo de 40% ao contrário da *Sustainalytics*.

Analisando a Figura 4.1 e 4.2, as variáveis do eixo do x, medem as atividades sustentáveis de um conjunto de empresas em relação à classificação por continente, surgindo como um indicador representativo do desenvolvimento sustentável. Verifica-se que a Europa apresenta valores ESG mais elevados e concentrados no 3º quartil, possuindo uma cauda mais longa à direita, interpretando que existe uma maior preocupação com questões de sustentabilidade na Europa do que no resto dos continentes, levando a que as empresas tenham uma classificação melhor, dado o seu esforço para combater e atingir as metas dos objetivos sustentáveis.

5. Modelação e interpretação de resultados

No presente capítulo serão aplicados e analisados os modelos de regressão utilizados para investigar a relação entre variáveis dependentes e independentes, ou seja, examinar como as variáveis relacionadas com o desempenho ambiental, social e de *governance* afetam a sustentabilidade geral das empresas. Essa relação é avaliada por meio dos coeficientes estimados que indicam a magnitude e a direção do impacto das variáveis independentes na sustentabilidade ESG, bem como a previsão ou explicação de um acontecimento com base nos dados observados.

A interpretação dos resultados da análise de regressão ESG é essencial para compreender os fatores que influenciam a sustentabilidade das empresas e identificar áreas de melhoria.

Além disso, a comparação entre os diferentes modelos de regressão permite avaliar a consistência e a robustez das relações encontradas. Por sua vez, a compreensão dessas relações pode auxiliar gestores, investidores e outras partes interessadas na tomada de decisões estratégicas e na promoção de práticas sustentáveis nas empresas.

De referir que, ao longo desta dissertação, é considerado um nível de significância de 0.10, por ser mais permissivo em relação à probabilidade de se cometer um erro do tipo I, isso significa que a barra para rejeitar a hipótese nula é mais baixa, o que pode levar a um maior número de resultados estatisticamente significativos. Corroborado a partir de um conjunto de aspetos na pesquisa como o tamanho da amostra, a magnitude do efeito, a precisão desejada, o contexto da área de estudo e as consequências dos erros de decisão, sendo melhor em termos teóricos.

Em suma, de seguida, serão examinados os resultados detalhados dos modelos de regressão e as suas implicações, a fim de fornecer uma visão mais aprofundada sobre os fatores que influenciam a sustentabilidade ESG, bem como as suas semelhanças e diferenças.

5.1 Modelo I: Fatores E, S, G e o Rating ESG

A análise, que será conduzida, terá como referência principal, o [anexo J](#). Baseado no facto de se possuir duas variáveis dependentes “R_ESG” e “Sus_ESG”, optou-se por realizar dois grupos de modelos de regressão.

$$Sus_{-ESG} = \beta_0 + \beta_1 E + \beta_2 S + \beta_3 G + \varepsilon \quad (13)$$

$$R_{-ESG} = \beta_0 + \beta_1 E + \beta_2 S + \beta_3 G + \varepsilon \quad (14)$$

Os resultados da regressão da *Sustainalytics* mostram que o modelo ajustado tem um coeficiente de determinação (R^2) de 0.971, o que significa que cerca de 97.1% da variabilidade da variável dependente "Sus_ESG" é explicada pelas variáveis independentes "Sus_E", "Sus_S", "Sus_G" incluídas no modelo. No entanto, no modelo da *Refinitiv*, é de apenas 30.1%, indicando que, no primeiro modelo, as variáveis independentes explicam melhor a variabilidade do ESG.

Os intervalos de confiança ([0.025, 0.975]) para os coeficientes indicam um nível de confiança de que os coeficientes reais estejam dentro dessa faixa, como exemplo, ter 95% de confiança de que o coeficiente "Sus_E" está entre 0.464 e 0.475. Neste caso ([anexo J](#)), os intervalos de confiança para todos os coeficientes não incluem zero, o que reforça a significância estatística dos coeficientes.

A análise dos coeficientes das variáveis independentes revela que todas elas têm um impacto positivo e estatisticamente significativos na variável dependente. Mais especificamente, para cada unidade de aumento em "Sus_E", espera-se um aumento de 0.4694 unidades em "Sus_ESG", mantendo as outras variáveis constantes. Da mesma forma, para cada unidade de aumento em "Sus_S" e "Sus_G", espera-se um aumento de 0.4597 e 0.2683 unidades em "Sus_ESG", respectivamente.

No caso apresentado, a estatística F, com um valor-p praticamente zero, mostra que as variáveis dos modelos como um todo são estatisticamente significativos, ou seja, pelo menos uma das variáveis independentes tem um efeito significativo sobre o ESG, rejeitando a hipótese nula que os coeficientes sejam todos zero.

No que se refere ao teste t, este avalia a significância individual de cada coeficiente estimado, verificando se o coeficiente é estatisticamente diferente de zero. No caso apresentado, indica que os coeficientes estimados são todos estatisticamente significativos, rejeitando a hipótese nula, podendo ter um impacto significativo na variável dependente.

É possível realçar um bom ajuste dos dados do modelo para a variável dependente ESG, com base no indicador *Log-Likelihood* (Log-verossimilhança), ou seja, quanto maior o valor do log-verossimilhança, melhor o ajuste do modelo. No caso apresentado, o valor do log-verossimilhança é -3375.9 na *Sustainalytics* e de -6478.6 na *Refinitiv*, logo melhor no primeiro modelo.

Os erros padrão dos coeficientes fornecem uma medida de precisão das estimativas, representando a variabilidade esperada nos coeficientes se o modelo fosse estimado várias vezes com diferentes amostras. Em síntese, quanto menor o valor do erro padrão, mais preciso é o coeficiente estimado, neste caso os erros padrão dos modelos não são expressivos.

Já no que se refere aos pressupostos dos resíduos, em relação à normalidade, estes não seguem uma distribuição normal, corroborado a partir do teste Jarque-Bera (JB) e Shapiro-

Wilk (anexo L), pois o valor-p é <0.10 , rejeitando a hipótese nula. Contudo, de acordo com o Teorema do Limite Central, a dimensão da amostra é suficientemente grande e, tendo em conta que a variação da população tem variância finita, considera-se satisfatória a aproximação à normal quando $n \geq 30$. Tal condição é satisfeita pelas amostras em análise, onde o valor de n é 1579.

De seguida, o gráfico de dispersão dos resíduos em relação aos valores ajustados (anexo K), no modelo da *Sustainalytics*, mostra que os pontos estão dispersos de forma uniforme ao redor da linha horizontal zero, o que sugere que a variância dos resíduos é constante em todos os níveis dos valores ajustados, indicando que os resíduos são homocedásticos. Adicionalmente, o pressuposto da ausência de relação entre resíduos e variáveis independentes também é atendido, sendo um indicador positivo de que o modelo está a capturar de maneira adequada, as relações entre as variáveis sem criar padrões nos resíduos que possam comprometer a interpretação ou validade das análises, fortalecendo a confiabilidade das conclusões derivadas do modelo de regressão linear múltipla. Além disso, o gráfico de dispersão dos resíduos em relação aos valores observados (anexo K), indica que a média dos resíduos está próxima de zero, o que, em média, os resíduos não estão a superestimar ou subestimar os valores observados.

Contrariamente, o gráfico do modelo da *Refinitiv*, indica que os resíduos (anexo K) têm um padrão de cone, funil ou leque ao longo da linha zero, indicando heteroscedasticidade, dependência das variáveis explicativas e dos resíduos e também uma média da distribuição de probabilidades dos resíduos diferente de zero, podendo esse padrão comprometer a validade das análises estatísticas e a interpretação dos resultados.

Segundo o teste *Durbin-Watson*, este é usado para detetar a presença de autocorrelação de 1ª ordem nos resíduos da regressão, o valor do teste varia entre 0 e 4, sendo que um valor próximo a 2 indica que não há autocorrelação nos resíduos. Neste caso, o valor do teste *Durbin-Watson* é de aproximadamente 1.983 e de 1.808, respetivamente, o que sugere a ausência de autocorrelação nos resíduos, o que é desejável num modelo de regressão, isso significa que os resíduos não estão correlacionados entre si.

As interpretações estatísticas dos parâmetros apresentados na regressão linear sugerem que cada parâmetro fornece informações importantes sobre o ajuste do modelo, da significância das variáveis independentes e da qualidade dos resíduos.

Deste modo, observa-se ainda uma acentuada assimetria negativa (anexo J) dos resíduos, indicando que a distribuição dos erros de previsão do modelo está inclinada para a esquerda, sugerindo que há uma tendência de subestimação dos valores da variável dependente em relação aos valores reais.

De seguida, a Curtose (anexo J), avalia a forma da distribuição dos resíduos do modelo, por outras palavras, mede o grau de achatamento dos resíduos em relação à distribuição

normal. Neste caso, o valor é de 542.559 na *Sustainalytics*, indicando uma curtose extremamente elevada, assinalando uma distribuição leptocúrtica mais concentrada e com caudas muito mais pesadas do que a distribuição normal, o que sugere uma elevada presença de valores extremos nos resíduos. Em relação à *Refinitiv*, não se verifica uma curtose tão elevada, contudo é superior a 3, apontando para o mesmo facto.

Pela análise acima, fica então demonstrado que cada fator E, S e G influencia o *score* final do *rating ESG*.

5.2 Variáveis *dummy*

Ao longo da análise dos dados, averiguou-se a natureza das variáveis do modelo concluindo que existiam variáveis categóricas (“*Region*”, “*PeerGroup*” e “*Index*”).

As variáveis *dummy*, são variáveis binárias criadas a partir de variáveis categóricas, onde cada categoria da variável original é representada por uma nova variável *dummy*. Essas variáveis assumem o valor 1 quando a categoria está presente e 0 quando não está, sendo usadas principalmente para representar informações categóricas em modelos estatísticos.

Numa primeira instância, optou-se por transformar a variável “*Region*” em *dummy*, com o intuito de analisar as interpretações do modelo por área geográfica, permitindo incluir informações categóricas no modelo, capturando diferenças entre categorias. De modo a evitar o problema da multicolinearidade, removeu-se a primeira coluna *dummy*, ou seja, a variável “*Region_Asia*”, definida como a categoria de referência.

Numa segunda instância, transformou-se a variável “*PeerGroup*” em *dummy*, com o objetivo de analisar as interpretações do modelo em relação aos diferentes setores das empresas. Apesar de, na amostra, estarem presentes 42 atividades distintas, ao longo da dissertação apenas se consideraram alguns setores, como o setor das Farmacêuticas, *Food Products* e *Oil & Gas Producers*.

Os setores selecionados foram escolhidos para representar uma variedade de indústrias com diferentes características e desafios em termos de sustentabilidade e responsabilidade corporativa. O setor Farmacêutico foi escolhido devido à sua importância na saúde humana e à necessidade de avaliar aspetos como ética na pesquisa e desenvolvimento de medicamentos, sustentabilidade e políticas de *governance* corporativa. O setor de *Food Products* foi incluído para examinar questões relacionadas com a segurança alimentar, práticas agrícolas sustentáveis, bem-estar animal e impactos ambientais na produção e consumo de alimentos. Já o setor de *Oil & Gas Producers* foi selecionado para analisar os desafios ESG associados à produção e exploração de combustíveis fósseis, como emissões de carbono, mudanças climáticas, gestão de resíduos e *governance* corporativa no setor de energia.

De referir que a variável “*Index*” não foi transformada em *dummy*, nem incluída nos modelos de regressão, justificada com base na lógica da multicolinearidade com a variável “*Region*”.

Essa multicolinearidade pode afetar negativamente a estabilidade dos resultados do modelo, tornando os coeficientes de regressão menos precisos, dificultando a interpretação dos efeitos das variáveis explicativas sobre a variável dependente.

Além disso, a inclusão de variáveis redundantes pode aumentar a complexidade do modelo sem adicionar benefícios significativos em termos de melhoria na qualidade das previsões ou interpretações dos resultados. Ou seja, se a variável “*Region*” já representa informações geográficas relevantes, como a localização das observações em diferentes regiões, a inclusão adicional da variável “*Index*”, que pode estar relacionada ao mesmo aspeto geográfico, pode ser redundante.

Por outro lado, a variável “*Index*” pode estar correlacionada com a variável “*Region*” uma vez que muitos índices de mercado de ações têm base regional. Se houver alta correlação entre essas variáveis, a inclusão de ambas pode dificultar a interpretação dos coeficientes de regressão e levar a estimativas imprecisas.

Com a utilização do *Teste Qui-quadrado*, foi possível verificar a associação entre as duas variáveis, ou seja, avaliar a independência entre as variáveis categóricas e verificar se há uma relação significativa entre elas. Neste caso ([anexo M](#)) o valor- $p < 0.10$, rejeitando a hipótese nula de independência e deduzindo-se que há uma associação significativa entre as variáveis, portanto, a categoria ou nível de uma variável pode influenciar a categoria ou nível da outra.

Para quantificar a força da associação utilizou-se a interpretação do coeficiente *V de Cramer*, este varia entre 0 e 1, onde 0 indica nenhuma associação entre as variáveis categóricas e 1 indica uma associação perfeita. Neste caso ([anexo M](#)) verifica-se uma associação perfeita.

Assim, para evitar problemas de multicolinearidade e garantir uma análise mais confiável, é mais apropriado omitir a variável “*Index*” já que a variável “*Region*” captura informações relevantes sobre a localização geográfica das observações.

De seguida, foram adicionadas variáveis de fatores financeiros e setoriais aos modelos, constituindo assim modelos capazes de analisar, tratar, presumir e retirar elações das classificações ESG.

5.3 Modelo II: Fatores geográficos, financeiros, E, S, G e o Rating ESG.

Neste modelo pretende-se explicar o *rating ESG* tendo em conta indicadores financeiros e geográficos, conforme detalhado no [anexo N](#).

O objetivo da análise de regressão é prever as variáveis dependentes com base num certo conjunto de variáveis independentes, que incluem “*Sus_E*”, “*Sus_S*”, “*Sus_G*”, “*R_E*”,

"R_S", "R_G", "MarkCap", "BookValue", "DivYield", "Region_Australia", "Region_Europe" e "Region_North America".

$$\begin{aligned}
 Sus_{ESG} = & -13.7058 + 0.4718E + 0.4641S + 0.2719G - 0.0002BookValue + \\
 & 5.116e^{-14}MarkCap - 0.8543DivYield - 0.1002Region_Australia - 0.2765Region_Europe \\
 & - 0.0929Region_North\ America + \varepsilon
 \end{aligned}
 \tag{15}$$

$$\begin{aligned}
 R_{ESG} = & 25.5773 + 0.1653E + 0.1395S + 0.2518G - 0.0001BookValue + \\
 & 3.598e^{-13}MarkCap + 11.9389DivYield + 4.4381Region_Australia + \\
 & 5.8178Region_Europe + 1.8336Region_North\ America + \varepsilon
 \end{aligned}
 \tag{16}$$

Com base no [anexo N](#), averigua-se que ambos os modelos i) são ajustados usando o método de regressão OLS; ii) têm uma variável dependente relacionada ao ESG e iii) usam variáveis independentes para prever essa variável.

No entanto, o R^2 associado ao modelo de regressão da *Sustainalytics*, indica que aproximadamente 98.1% da variabilidade em "Sus_ESG" é explicada pelas variáveis independentes, já no modelo da *Refinitiv*, o R^2 é mais baixo, indicando que apenas cerca de 31.6% da variabilidade em "R_ESG" é explicada pelas variáveis independentes. Concluindo-se que no geral o primeiro modelo tem um melhor ajuste, indicando que as variáveis independentes explicam melhor a variabilidade do ESG.

Os coeficientes da constante são de -13.7058 e de 25.5773 no primeiro e segundo modelo, respetivamente, o que remete para o valor médio esperado de ESG quando todas as outras variáveis independentes são iguais a zero e a variável categórica "Region" é igual a Ásia.

No modelo da *Sustainalytics*, todas as variáveis independentes possuem coeficientes estatisticamente significativos (valor-p < 0.10), ou seja, existe uma relação significativa entre as variáveis consideradas, rejeitando a hipótese nula, exceto nas variáveis "Region_Australia", "Region_North America" e "DivYield". Este facto sugere que não são estatisticamente significativas (valor-p > 0.10), propondo que não há evidências estatísticas suficientes para afirmar que essas variáveis têm um impacto significativo na variável dependente ("Sus_ESG").

Por conseguinte, à semelhança do modelo anterior, o da *Refinitiv*, segue a mesma lógica, à exceção da variável "Region_Australia", rejeitando a hipótese nula e da variável "MarkCap", não rejeitando a hipótese nula. Esta diferença na significância estatística da variável "Region_Australia" e "MarkCap" nos dois modelos pode ser atribuída a uma série de fatores, incluindo a relação das variáveis com a variável dependente e a interação com outras

variáveis independentes no modelo. Pois, no segundo modelo, a variável "*MarkCap*" possui um valor-p > 0.10, indicando que não há evidências suficientes para afirmar que o coeficiente estimado é diferente de zero ao nível de significância usual, sugerindo que a inclusão dessa variável no modelo não contribui de forma estatisticamente significativa para explicar a variabilidade na variável dependente "*R_ESG*".

De referir que, apesar de serem apresentados coeficientes negativos no modelo da *Sustainalytics*, pode dizer-se que não há evidências estatísticas suficientes para afirmar que há uma relação inversa significativa entre a variável independente "*Region_Australia*", "*Region_North America*" e "*DivYield*" e a variável dependente "*Sus_ESG*".

No entanto, pode-se concluir que o coeficiente negativo da variável "*Region_Europe*" denota que as empresas localizadas na Europa podem ter uma pontuação de sustentabilidade ESG menor em comparação com a região de referência, mas não se pode afirmar que é necessariamente mais baixa do que as empresas da Ásia.

Por outro lado, o modelo da *Refinitiv* apresenta coeficientes positivos para as variáveis "*Region_Australia*" e "*Region_Europe*", o que assinala que as empresas localizadas nessas regiões tendem a ter um desempenho de sustentabilidade ESG superior em comparação com as empresas da região de referência.

Tendo em conta ambos os modelos, a diferença na significância estatística entre as variáveis "*Region_North America*" e "*Region_Europe*" pode ser explicada por diversos fatores tais como regulamentações distintas, sensibilidade à sustentabilidade, cultura empresarial e composição setorial. A Europa tem sido historicamente mais rigorosa em termos de regulamentação ambiental e social, sendo que as empresas podem estar mais sujeitas a padrões e expectativas específicas de ESG, o que pode aumentar a relevância da variável nos modelos. Além disso, a pressão pública e as prioridades de investidores em relação a ESG também podem diferir entre as duas regiões. Contudo, a falta de significância estatística da variável "*Region_North America*", não implica que as práticas de ESG sejam menos importantes nos EUA, mas sim que outros fatores podem desempenhar um papel mais proeminente na determinação do *rating ESG* das empresas nessa região.

Por conseguinte, pode-se inferir que, em ambos os modelos, o coeficiente da variável "*BookValue*" é negativo e estatisticamente significativo, o que cita que um aumento no valor contabilístico está associado a uma diminuição no *rating ESG*. Essa relação pode ser justificada, por exemplo, pelo fato de que um valor mais elevado de "*BookValue*" poder indicar a presença de ativos físicos substanciais, tais como propriedades ou infraestruturas, que podem ter um impacto negativo no desempenho ambiental das empresas. Além disso, também pode refletir um maior nível de endividamento, o que pode dificultar para as empresas o investimento em práticas sustentáveis.

Em relação aos pressupostos dos resíduos dos modelos de regressão, ambos verificam as mesmas conclusões referidas anteriormente, na secção 5.1, em que o primeiro modelo cumpre todos os pressupostos, enquanto o segundo modelo viola nomeadamente, o pressuposto da variância, da ausência de relação entre resíduos e variáveis independentes e da média 0, afetando a precisão dos intervalos de confiança e os testes de significância dos coeficientes.

Tendo em conta o referido anteriormente e levando em conta o critério AIC e BIC, o modelo da *Sustainalytics* possui maior qualidade de ajuste dos dados.

5.4 Modelo III: Fatores financeiros, E, S, G e o Rating ESG

Neste modelo, apenas se tem em consideração fatores financeiros, como o retorno dos dividendos, o valor contabilístico e valor de mercado, representados pelas variáveis “*DivYield*”, “*BookValue*” e “*MarkCap*”, respetivamente, do conjunto de empresas dos 20 índices da bolsa de valores, com o intuito de analisar o desempenho em termos de poder explicativo no *rating ESG* dos modelos, com um menor número de variáveis independentes, tendo em conta o [anexo O](#).

$$Sus_{-ESG} = -13.5791 + 0.4712E + 0.4637S + 0.2694G - 0.0002BookValue - 0.9636DivYield + \varepsilon \quad (17)$$

$$R_{-ESG} = 25.7523 + 0.1554E + 0.1652S + 0.2742G - 0.0001BookValue + 14.5548DivYield + \varepsilon \quad (18)$$

Primeiramente, verifica-se que, em ambos os modelos, os fatores E, S e G influenciam o *score* final do *rating ESG*. Considerando, por exemplo, duas empresas do setor de automóveis, a Empresa A e a Empresa B, em que ambas investem em tecnologias mais limpas e eficientes para reduzir as suas pegadas de carbono e adotam práticas mais sustentáveis nas suas operações, no entanto apresentam valores contabilísticos dispares, um valor de 500 M\$ e de 2 bi \$, respetivamente. Este facto sugere que a Empresa B, com um valor contabilístico mais elevado, pode ter uma pontuação ESG ligeiramente menor em comparação com a Empresa A. Isso pode levantar a hipótese de que a Empresa B pode estar a investir mais em expansão e aquisições, o que pode estar relacionado com práticas menos sustentáveis, resultando num *score ESG* marginalmente inferior. No entanto, a diferença na pontuação ESG é mínima uma vez que a redução estimada é de apenas 0.0001 ou 0.0002 na pontuação ESG da Empresa B em relação à Empresa A. Assim à semelhança do modelo I, o coeficiente da variável “*BookValue*” é negativo e estatisticamente significativo.

Por sua vez, consideram-se duas empresas no setor de tecnologia, a Empresa X com um “*DivYield*” de 2% e a Empresa Y com um “*DivYield*” de 5%. Com base no coeficiente positivo do “*DivYield*”, esperar-se-ia que a Empresa Y tivesse um *score ESG* ligeiramente superior ao da Empresa X. Isso pode sugerir que a Empresa Y está a distribuir mais lucros aos acionistas na forma de dividendos e pode ser visto como uma prática financeira positiva, mas não necessariamente relacionada às métricas ESG. No entanto, como o coeficiente para a variável “*DivYield*” não é estatisticamente significativo, não se pode afirmar com confiança que o rendimento de dividendos tem um impacto significativo na pontuação ESG dessas empresas.

Observa-se que os coeficientes estimados diferem dos dois modelos apresentados na secção anterior, o que pode ser atribuído às diferenças nas variáveis independentes incluídas em cada modelo. Essa discrepância ocorre devido ao facto de que cada variável independente possuir uma relação única com a variável dependente e a inclusão ou exclusão de determinadas variáveis pode resultar em estimativas diferentes dos coeficientes.

Apesar da redução no número de variáveis independentes em comparação com os modelos anteriores, pode-se conferir uma abordagem mais focada na seleção das variáveis financeiras mais relevantes para explicar a variabilidade do ESG, conduzindo a uma melhor interpretação e compreensão dos coeficientes estimados no modelo em questão. Ao considerar esse modelo, é importante ressaltar que, embora a inclusão de menos variáveis possa reduzir a complexidade do modelo, também pode resultar numa explicação menos abrangente da variabilidade do ESG. No entanto, se o objetivo principal é identificar as variáveis financeiras mais influentes e significativas para a sustentabilidade ESG, este modelo pode ser preferível.

De referir que são retiradas as mesmas ilações da secção 5.3, em que o modelo da *Sustainalytics* aparenta ter um desempenho melhor em termos de explicação da variabilidade do ESG e pressupostos dos resíduos.

5.5 Modelo IV: Fatores E, S, G, financeiros, geográficos, setoriais e o Rating ESG

Neste modelo pretende-se estudar o *rating ESG*, não só tendo em conta fatores E, S, G, financeiros e geográficos, mas também analisar se o setor de atividade das empresas tem efeito na previsão do ESG, conforme apresentado no [\(anexo P\)](#).

$$\begin{aligned}
 Sus_{-ESG} = & -13.6735 + 0.4717E + 0.4638S + 0.2719G - 0.0002BookValue + \\
 & 5.165e^{-14}MarkCap - 0.8428DivYield - 0.1004Region_Australia - 0.2732Region_Europe - \\
 & 0.0918Region_North\ America + PeerGroup + \varepsilon
 \end{aligned}
 \tag{19}$$

$$R_{ESG} = 25.5830 + 0.1662E + 0.1380S + 0.2502G - 0.0001BookValue + 3.573e^{-13} MarkCap + 11.5922DivYield + 4.5095Region_Australia + 5.8317Region_Europe + 1.8517Region_North\ America + PeerGroup + \varepsilon \quad (20)$$

Com a análise do [anexo P](#), infere-se que o modelo da *Sustainalytics* é, mais uma vez, o mais forte, porém a inserção das variáveis independentes direcionadas ao setor de atividade não permite demonstrar uma relação estatisticamente significativa com a variável dependente, no contexto dos modelos em questão. Isto sugere que, uma vez que outras variáveis relacionadas a aspectos ESG (como "Sus_E", "Sus_S", "Sus_G", entre outras) são consideradas no modelo, a contribuição específica do setor de atividade não pode ser distinguida de outros fatores já incluídos no modelo, no entanto não se pode deduzir que o setor de atividade das empresas não influencia o *rating ESG*, pois essa afirmação é limitada ao objetivo específico do modelo e dos dados disponíveis, mas sim que não determina diretamente o seu *rating ESG*. O mesmo se verifica no modelo da *Refinitiv*, em que o valor-p > 0.10, não rejeitando a hipótese nula.

A título de exemplo, é possível considerar que empresas do setor de *Oil & Gas* possuam classificações ESG altas, indicando um bom desempenho em aspectos ambientais, sociais e de *governance*, ao mesmo tempo em que são consideradas as mais poluentes. Essa aparente contradição pode ser explicada por diferentes fatores.

Primeiramente, é importante considerar os critérios de avaliação utilizados para atribuir as classificações ESG, pois estes podem variar e enfatizar diferentes aspectos do desempenho das empresas, assim, uma empresa desse setor pode ser avaliada positivamente em aspectos como a *governance* corporativa e transparência em relatórios, compensando as suas emissões de poluentes. Além disso, muitas empresas desse setor implementam medidas de compensação e mitigação para reduzir o seu impacto ambiental, investindo em tecnologias mais limpas, promovendo iniciativas de eficiência energética, adotando o uso de energias renováveis, compensando as suas emissões de carbono por meio de projetos de reflorestamento ou compra de créditos de carbono.

A adoção dessas práticas pode depender de uma combinação complexa de estratégias de negócios, regulamentações específicas do setor e fatores contextuais. Para exemplificar: uma empresa de Alimentos pode estar sujeita a regulamentações rigorosas relacionadas à rotulagem de ingredientes e práticas agrícolas sustentáveis, influenciando as suas práticas ESG, já uma empresa de Finanças, pode priorizar a transparência nas suas políticas de investimento responsável, visando impactos positivos na sociedade e no meio ambiente. Assentando a importância de uma abordagem flexível e interdisciplinar ao avaliar a influência do *rating ESG*, reconhecendo a complexidade das relações entre fatores financeiros e considerações sustentáveis num ambiente empresarial diversificado.

De referir que são retiradas as mesmas ilações da secção 5.3, em que o modelo da *Sustainalytics* aparenta um desempenho superior em termos de explicação da variabilidade do ESG, pressupostos dos resíduos e em termos da significância estatística.

5.6 Testes estatísticos

Para avaliar a significância estatística das variáveis independentes em relação às variáveis dependentes, foram realizados o teste t e o teste ANOVA, ambos testes estatísticos, utilizados para avaliar se existem diferenças estatisticamente significativas entre as médias de dois grupos.

O teste ANOVA (Análise de Variância) avalia a variabilidade entre os grupos em relação à variabilidade dentro de cada grupo. O teste calcula uma estatística F e um valor-p que indicam se existem diferenças estatisticamente significativas entre as médias dos grupos. Neste estudo, com base no [anexo Q](#), em todos os casos se rejeita a hipótese nula, o que significa que há pelo menos uma variável explicativa que contribui para justificar o modelo.

O teste t de *Student* calcula a diferença entre as médias das amostras em relação à variabilidade dos dados dentro de cada grupo, expressa pelo desvio padrão. O resultado do teste é uma estatística t e um valor-p que indicam se as diferenças entre as médias são estatisticamente significativas. De referir que apenas se torna necessário realizar este teste quando se rejeita a hipótese nula no teste F, nesta situação rejeita-se a hipótese nula em todos os casos, o que significa que todas as variáveis explicativas consideradas contribuem para justificar o modelo, conforme o [anexo Q](#).

Por exemplo, ao comparar as variáveis "Sus_E" e "Sus_ESG", tanto o teste t de *Student* quanto o teste ANOVA resultaram em valores-p de quase zero, indicando uma diferença estatisticamente significativa entre as médias dessas variáveis. O mesmo ocorre nas comparações entre "Sus_S" e "Sus_ESG", "Sus_G" e "Sus_ESG", bem como entre "DivYield" e "Sus_ESG", "BookValue" e "Sus_ESG", "MarkCap" e "Sus_ESG", "Region_Australia" e "Sus_ESG", "Region_Europe" e "Sus_ESG", "Region_North America" e "Sus_ESG", "PeerGroup_Chemicals" e "Sus_ESG", "PeerGroup_Oil & Gas Producers" e "Sus_ESG", "PeerGroup_Pharmaceuticals" e "Sus_ESG", "PeerGroup_Precious Metals" e "Sus_ESG", e "PeerGroup_Refiners & Pipelines" e "Sus_ESG".

Portanto, pode inferir-se que essas variáveis têm influência significativa no *rating ESG*, quando analisado individualmente, sugerindo que diferentes setores podem ter diferentes desempenhos em termos de critérios ESG, contribuindo assim para a explicação da variabilidade dos dados.

5.7 Modelo V: Seleção das variáveis mais significativas com base na significância individual

A seleção de variáveis é um processo crucial na análise de dados que visa identificar e escolher as características mais relevantes e informativas para a construção de um modelo preditivo. Em muitos cenários, as bases de dados podem conter diversas variáveis, algumas das quais podem não ter impacto significativo na variável de interesse.

A inclusão dessas variáveis não relevantes pode levar a modelos excessivamente complexos com risco de *overfitting*. Para superar esse desafio, diferentes técnicas de seleção de variáveis são aplicadas, permitindo a obtenção de modelos mais simples e precisos.

O método *SelectKBest* da biblioteca *scikit-learn* do *Python*, em conjunto com a regressão, é uma das técnicas utilizadas para selecionar as variáveis mais relevantes. Nesse método, é aplicado um teste estatístico como o teste F de regressão, para medir a correlação linear entre cada variável independente e a variável dependente. Com base nos resultados desse teste, o *SelectKBest* escolhe as "k" melhores variáveis com as maiores correlações, onde "k" é um valor pré-definido, mais especificamente que avalia a capacidade preditiva individual de cada variável em relação à variável dependente, sem levar em conta as outras variáveis do modelo. Neste caso e conforme o [anexo R](#), é definido "k=5", onde pode ser consultada a escolha das 5 melhores variáveis.

Primeiramente no modelo da *Sustainalytics* ([anexo R](#)), foram selecionadas as seguintes variáveis: "Sus_E", "Sus_S", "Sus_G", "Region_Europe" e "PeerGroup_Oil & Gas Producers".

Em resumo, o primeiro modelo de regressão tem um melhor ajuste aos dados, explicando aproximadamente 97.1% da variabilidade em "Sus_ESG". As variáveis "Sus_E", "Sus_S" e "Sus_G" são estatisticamente significativas e têm uma forte relação linear positiva com "Sus_ESG", já a variável "Region_Europe" torna-se estatisticamente significativa pois o valor-p < 0.10, o que indica que há uma relação significativa entre a região da Europa e a pontuação ESG, no entanto, o coeficiente negativo sugere que a pontuação ESG na Europa é menor em comparação com a Ásia. É importante lembrar que, embora seja estatisticamente significativa, a significância é menor em relação às outras variáveis.

Por último, a variável "PeerGroup_Oil & Gas Producers" não é estatisticamente significativa, pois o valor-p > 0.10 indicando que, isoladamente, não tem um impacto estatisticamente significativo na variável dependente "Sus_ESG". Surpreendentemente, a variável categórica foi selecionada como uma das cinco melhores variáveis neste modelo específico, indicando uma possível relação estatisticamente significativa com a variável dependente. Todavia, isso é justificado com base no modelo *SelectKBest*, pois este pode selecionar as melhores variáveis com base na sua capacidade preditiva individual e em como elas contribuem para explicar a variabilidade na variável dependente. No entanto, quando

essas variáveis são consideradas em conjunto num modelo de regressão linear, o impacto estatístico de cada variável pode mudar devido às interações entre elas.

Em segundo lugar, no modelo da *Refinitiv* (anexo R), foram selecionadas as seguintes variáveis: “R_E”, “R_S”, “R_G”, “*Region_Europe*” e “*Region_North America*”. Os resultados indicam que as variáveis “R_E”, “R_S”, “R_G”, são estatisticamente significativas e têm uma forte relação linear com o *rating ESG*, já a variável “*Region_Europe*” apresenta um impacto significativo no ESG, rejeitando a hipótese nula, ainda que em menor escala, ou seja, empresas localizadas na região da Europa tendem a ter uma pontuação ESG mais alta do que as empresas localizadas na Ásia. Finalmente, a variável “*Region_North America*”, apesar de selecionada como uma das 5 melhores variáveis pelo método *SelectKBest*, não é estatisticamente significativa no modelo de regressão, o que sugere que, quando se considera todas as variáveis em conjunto, ela não tem um efeito significativo na variação da pontuação “R_ESG” além daquelas que já foram incluídas no modelo.

Para concluir, são retiradas as mesmas ilações da secção 5.3 e 5.4, em que o modelo da *Sustainalytics* aparenta um desempenho superior em termos de explicação da variabilidade do ESG, pressupostos dos resíduos e em termos da significância estatística.

5.8 Multicolinearidade

Com base nos resultados do cálculo do VIF (*Variance Inflation Factor*), pode-se afirmar que os dados utilizados nos modelos de regressão não apresentam multicolinearidade significativa entre as variáveis independentes, de acordo com o anexo S.

A maioria das variáveis possui VIFs próximos a 1, indicando que elas não estão altamente correlacionadas entre si. Essa constatação é importante pois reduz o risco de problemas de instabilidade e enviesamento nos resultados do modelo. No entanto, é necessário estar ciente da correlação moderada entre as variáveis “*Region_Europe*” e “*Region_North America*” com valores de VIF de 2.46 e 2.69, respetivamente. Apesar disso, os VIFs dessas variáveis ainda estão abaixo do limiar de 5, considerado aceitável. Portanto, pode-se considerar que os dados estão adequados para a construção dos modelos de regressão pois a multicolinearidade não apresenta uma preocupação significativa.

5.9 Matriz de correlação

Construiu-se uma matriz de correlação para quantificar a relação linear entre as variáveis independentes e as variáveis dependentes.

Através do mapa de calor (Figura 5.1), podem-se identificar padrões de correlação entre as variáveis. As cores mais quentes (como vermelho) indicam uma correlação positiva mais forte, cores mais frias (como azul) indicam uma correlação negativa mais forte e cores neutras (como branco) indicam uma correlação próxima de zero ou ausência de correlação.

Primeiramente, no que diz respeito às métricas ESG e Pontuação ESG, observa-se que há uma correlação positiva forte entre as variáveis independentes ESG (“Sus_E”, “Sus_S”, “Sus_G”, “R_E”, “R_S”, “R_G”) e a variável dependente (“Sus_ESG”, “R_ESG”, respectivamente), o que revela que as empresas com bom desempenho em termos ambientais, sociais e de *governance* também tendem a ter uma pontuação ESG mais elevada.

Por sua vez, as métricas ESG também mostram correlações positivas significativas entre si como no caso das variáveis “Sus_S” e “Sus_G”, bem como nas variáveis “R_E” e “R_S”, “R_E” e “R_G” e “R_S” e “R_G”, o que inculca que empresas que se destacam num dos aspetos ESG tendem a ter bom desempenho noutros aspetos ESG, contudo em menor escala comparativamente com o ponto anterior. Adicionalmente, quando “Sus_E” aumenta, “Sus_S” e “Sus_G” tende a diminuir e vice-versa.

De seguida, a nível das métricas Financeiras, como as variáveis “*DivYield*”, “*BookValue*” e “*MarkCap*”, não apresentam correlações fortes com as métricas ESG, apontando que o desempenho financeiro das empresas não tem uma relação direta significativa com os seus aspetos ESG.

Por conseguinte, a localização Geográfica das variáveis categóricas (“*Region_Australia*”, “*Region_Europe*”, “*Region_North America*”), mostra correlações fracas com as métricas ESG e financeiras, insinuando que a localização geográfica pode não ser um fator decisivo para determinar o desempenho ESG e financeiro das empresas.

De igual modo, as métricas ESG e financeiras não apresentam correlações significativas com os diferentes setores da indústria (“*PeerGroup_Chemicals*”, “*PeerGroup_Oil & Gas Producers*”, entre outros), enunciando que o desempenho ESG e financeiro pode variar, independentemente do setor de atividade da empresa. À exceção da relação da variável “Sus_E” e “*PeerGroup_Oil & Gas Producers*”, “Sus_S” e “*PeerGroup_Pharmaceuticals*” e “Sus_ESG” e “*PeerGroup_Oil & Gas Producers*”, na qual possuem uma correlação significativa negativa de -0.29, -0.26 e -0.25, respetivamente. A título de exemplo, uma empresa que se dedica a produzir energia a partir de fontes sustentáveis, como solar e eólica, com o objetivo de reduzir a dependência de combustíveis fósseis e minimizar as emissões de gases de efeito estufa, pode apresentar um índice de sustentabilidade ambiental (“Sus_E”) relativamente baixo devido ao histórico da indústria de combustíveis fósseis, ao qual já esteve associada antes da transição para energias renováveis. Em seguida, por exemplo, uma Farmacêutica apresenta um índice de sustentabilidade social (“Sus_S”) relativamente baixo pois, apesar do seu papel importante na área da saúde, a empresa pode receber uma pontuação de sustentabilidade social mais baixa devido a críticas ou controvérsias em relação a práticas de precificação e distribuição dos produtos.

Além disso, a ausência de correlações significativas entre as variáveis dependentes, sugere que não há problemas de multicolinearidade no conjunto de dados, o que é importante para a precisão dos modelos estatísticos utilizados.

No entanto, se se observarem especificamente os coeficientes de correlação da *Sustainalytics* com a *Refinitiv*, como esperado, apresentam-se na maioria negativos. Assim, de forma geral, é possível verificar que há uma relação inversa (divergência) entre as classificações ESG da *Sustainalytics* e da *Refinitiv*, conforme já detalhado na revisão da literatura.

Mais importante ainda, isso mostra aos investidores a importância de comparar duas classificações ESG diferentes quando chega a hora de investir, simplesmente porque elas medem coisas diferentes, tendo em vista que não há um referencial consensual nessa questão.

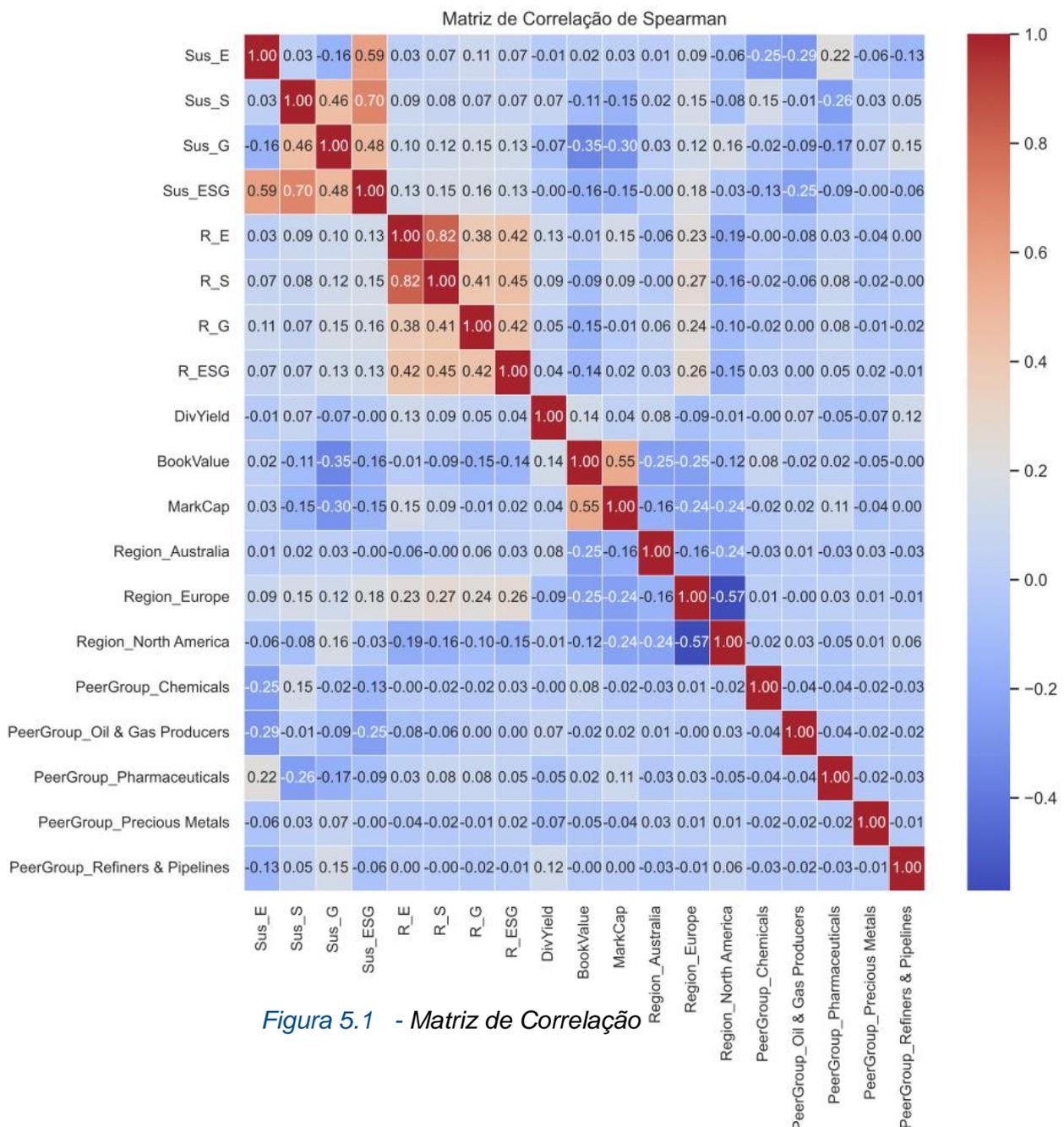


Figura 5.1 - Matriz de Correlação

6. Avaliação e previsão

A divisão dos dados num conjunto de treino e teste é uma prática fundamental na análise estatística, esse processo visa avaliar a capacidade de generalização do modelo e garantir a sua eficácia em dados não vistos previamente.

Ao dividir os dados em dois conjuntos independentes, o conjunto de treino é utilizado para ajustar os parâmetros do modelo enquanto o conjunto de teste é reservado para avaliar o desempenho do modelo em dados não utilizados durante o treino. Isso é essencial para evitar o *overfitting*, um cenário onde o modelo se adapta excessivamente aos dados de treino, perdendo a capacidade de generalizar para novos dados.

A proporção do conjunto de treino e teste é de 80% dos dados para treino e 20% para teste.

A técnica utilizada nesta dissertação é a validação cruzada *K-Fold*, que é uma abordagem de validação de modelos que visa avaliar o desempenho do modelo em dados não vistos previamente, com o objetivo de dividir o conjunto de dados em "k" partes ou *folds*, sendo que em cada iteração, um *fold* é utilizado como conjunto de teste e os restantes como conjunto de treino. Isso é repetido "k" vezes, de forma que cada *fold* seja usado exatamente uma vez como conjunto de teste.

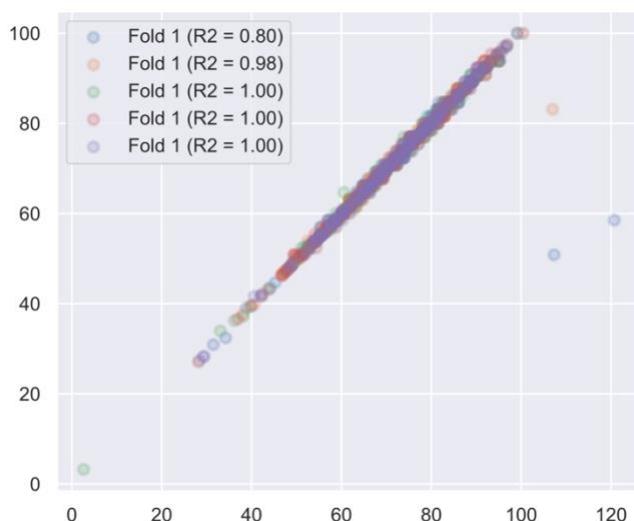


Figura 6.1 - Validação cruzada K-Fold da Sustainalytics

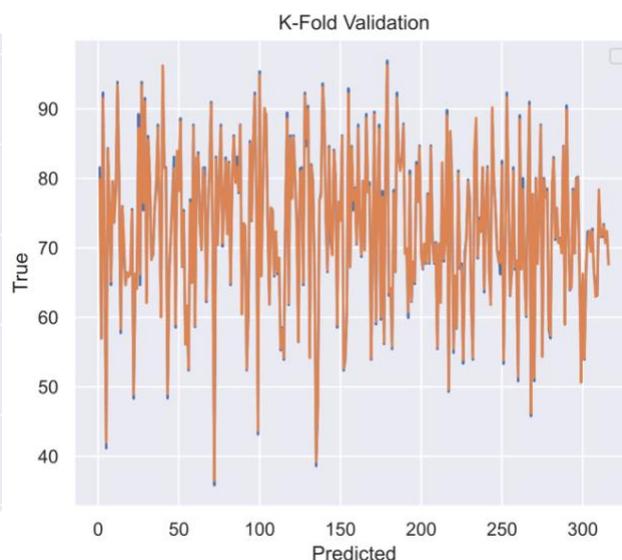


Figura 6.2- Valores previstos vs valores reais da Sustainalytics

Verifica-se que a figura 6.1 da Sustainalytics, no contexto da validação cruzada *K-Fold*, cada ponto no gráfico representa uma observação no conjunto de teste, enquanto a linha de regressão linear positiva representa a relação ajustada entre a variável alvo original e os valores preditos pelo modelo de regressão, o que significa que à medida que o modelo prevê valores mais altos para a variável dependente, os valores reais também tendem a ser mais altos.

Ao analisar os valores de R^2 para cada *fold*, pode-se interpretar o desempenho do modelo em cada iteração da validação cruzada. Quando o R^2 é próximo a 1, indica um ajuste quase perfeito do modelo aos dados, o que sugere que o modelo é capaz de fazer boas previsões em diferentes conjuntos de dados de treino e teste.

Reforçando a análise e tendo em conta a [figura 6.2](#), o modelo é capaz de prever quase exatamente os valores reais (linha azul) para a variável dependente, apresentando um bom desempenho geral, uma elevada capacidade de explicar a variabilidade da variável dependente, concluindo-se que quanto mais os pontos previstos (linha laranja) se aproximarem da linha real, melhor o modelo estará ajustando aos dados e a fazer previsões precisas.

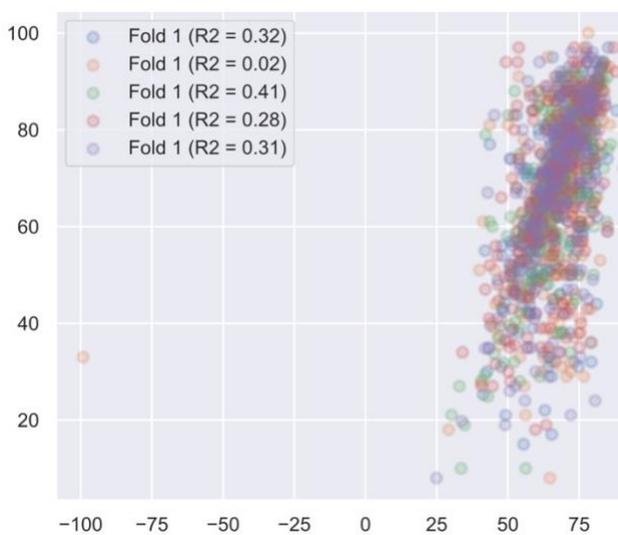


Figura 6.3 - Validação cruzada K-Fold da Refinitiv

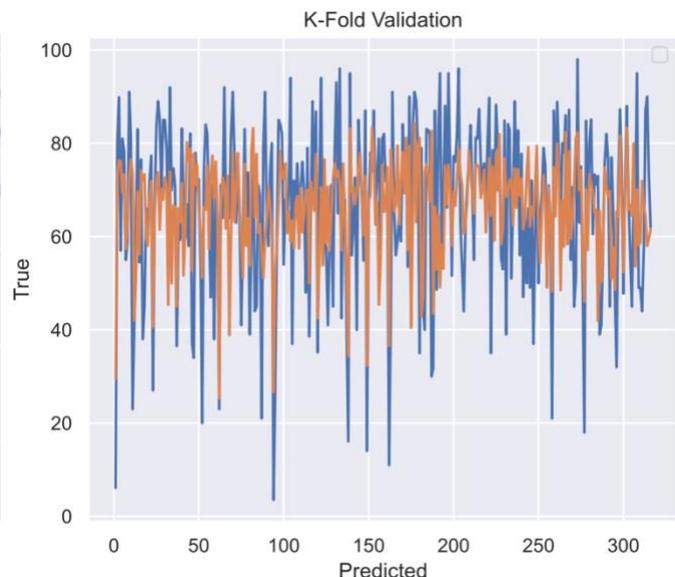


Figura 6.4 - Valores previstos vs valores reais da Refinitiv

Em relação à [figura 6.3](#) da *Refinitiv*, mostra que os pontos estão mais dispersos em relação à linha de regressão, indicando que a relação entre as variáveis independentes e a variável dependente é mais fraca e menos linear em comparação com o modelo anterior.

A dispersão dos pontos em relação à linha de regressão indica que o modelo pode não ser capaz de fazer previsões tão precisas para a variável “R_ESG”, devido à natureza mais complexa e menos linear da relação entre as variáveis.

No que toca aos valores de R^2 são inferiores, variando entre 0.02 e 0.41, para o *fold* onde está presente o *outlier* que pode ser visualizado na [figura 6.3](#), o que indica que o modelo tem dificuldade em explicar a variabilidade da variável “R_ESG” e não se ajusta tão bem aos dados.

No que diz respeito à [figura 6.4](#), observa-se que a previsão gerada pelo modelo de regressão linear mostra falta de precisão e acurácia, relativamente à variável “R_ESG”, sendo isto, evidenciado pelo facto da linha laranja estar muito mais compacta e menor, em relação à linha azul do gráfico, sugerindo que o modelo mostra dificuldade em capturar a variabilidade dos dados reais, acabando por gerar previsões muito similares para diferentes níveis da variável dependente.

Em suma, a comparação entre os dois modelos indica que o modelo da *Sustainalytics* apresenta um desempenho superior ao modelo da *Refinitiv* em relação à precisão das previsões e à explicação da variabilidade da variável dependente.

7. Considerações finais

A conclusão deste estudo é fruto de uma investigação minuciosa da relação entre o *rating ESG* e os métodos utilizados e desenvolvidos pela matemática financeira. Esta é enriquecida pela integração abrangente de áreas como a gestão de informação, inteligência artificial, programação e a análise e qualidade de dados. Cada secção visa o culminar de reafirmar os objetivos traçados, destacar as descobertas notáveis, discutir as contribuições para o campo do investimento sustentável e responsável e validar as limitações inerentes ao estudo, bem como as potenciais direções para futuras investigações.

Primeiramente, valida-se o argumento que contribuiu para a compreensão cuidada das interações complexas entre fatores e o seu impacto no desempenho das empresas e nas decisões de investimento sustentável. Reiterando, enquanto um *rating* tradicional avalia a capacidade de crédito de uma entidade, um *rating ESG* abrange a avaliação do desempenho de uma empresa em termos de sustentabilidade ambiental, responsabilidade social e práticas de *governance* corporativa, indo além dos indicadores de risco de crédito para considerar um espectro mais amplo de critérios que refletem o impacto social e ambiental de uma empresa, bem como a sua *governance* interna.

Em segundo lugar, os objetivos iniciais neste estudo foram alcançados pois, ao avaliar os *ratings ESG* dos 20 índices da bolsa de valores em diferentes setores e partes do mundo e por sua vez da influência dos *ratings* no desempenho das empresas, respondeu-se com complexidade à questão primordial que norteou esta investigação, “*Até que ponto um rating ESG mais elevado influencia o desempenho das empresas?*” proporcionando uma visão abrangente da influencia intrincada na rede de fatores que moldam a avaliação ESG. Assim, embora haja indícios de uma associação positiva entre boas performances ESG e *rating ESG*, essa relação pode ser influenciada por outros fatores, destacando a complexidade das interações entre as variáveis. Isso ressalta a importância de continuar a investigação e compreender as relações entre as dimensões ESG e o desempenho corporativo de maneira mais detalhada.

Com base nos resultados obtidos através do modelo da *Sustainalytics*, identificou-se que as variáveis estatisticamente significativas, “*Sus_E*”, “*Sus_S*”, “*Sus_G*”, “*MarkCap*”, “*BookValue*” e “*Region_Europe*” poderiam influenciar o *rating ESG*, observando padrões distintos de influência. As variáveis ambientais, sociais e de *governance* influenciam positivamente o desempenho ESG. Quando se analisa a variável “*MarkCap*” com um coeficiente negativo, isso pode ser interpretado como um reflexo em que as empresas com um valor de mercado elevado possam estar suscetíveis a certas pressões competitivas e aumento da procura por parte de investidores, o que pode distanciar a atenção e os recursos para práticas mais sustentáveis a longo prazo.

Em relação à variável “*Book Value*”, o coeficiente negativo sugere que as empresas com ativos contábilísticos mais elevados possam vir a enfrentar maiores desafios na implementação de práticas sustentáveis, possivelmente devido a compromissos financeiros pré-existentes ou à alocação de recursos para a manutenção e desenvolvimento de ativos tangíveis. Por fim, a variável “*Region Europe*”, com coeficiente negativo, pode indicar que as empresas na região europeia enfrentam desafios específicos ou podem ter regulamentações mais rigorosas em relação às práticas ESG, pois têm registros de uma longa história de ativismo ambiental e social, o que pode resultar em maior consciencialização e pressão sobre as empresas em relação às práticas de ESG, contribuindo para a relevância da variável nos modelos.

No âmbito do modelo da *Refinitiv*, observou-se que as variáveis “*R_E*”, “*R_S*”, “*R_G*”, “*BookValue*”, “*Region_Australia*” e “*Region_Europe*” poderiam influenciar o *rating ESG*. As variáveis ambientais, sociais e de *governance* possuem um coeficiente positivo, à semelhança do modelo anterior. A variável “*BookValue*” destaca-se novamente com um coeficiente negativo, apontando para desafios semelhantes enfrentados por empresas com ativos contábilísticos mais substanciais, independentemente da região geográfica.

Por sua vez, a variável “*Region Australia*” com coeficiente positivo, pode indicar que as empresas nesta região têm conseguido combinar práticas ESG mais sólidas com um bom desempenho de sustentabilidade, possivelmente beneficiando de regulamentações favoráveis e crescente consciencialização sobre a sustentabilidade, visto que é um país desenvolvido e faz parte do grupo de países com economias mais avançadas. Por último, a variável “*Region Europe*”, com coeficiente positivo, neste contexto pode ser interpretado como um contraste interessante com o modelo anterior propondo que, enquanto algumas empresas europeias enfrentam desafios nas suas práticas ESG, outras na mesma região têm conseguido superar essas dificuldades e demonstrar um desempenho sólido nas métricas ESG, destacando assim a complexidade das dinâmicas regionais e como as abordagens de sustentabilidade podem variar significativamente entre diferentes partes do mundo.

Consequentemente, a análise de regressões revelou resultados intrigantes, destacando variáveis que, embora tenham sido incorporadas nos modelos, não demonstraram ter relevância estatística em relação à influência no *rating ESG*, especificamente as variáveis, “*Region_North America*”, “*DivYield*” e os “*PeerGroup*”.

No caso da variável “*Region_North America*”, a sua falta de significância estatística sugere que a localização geográfica nos Estados Unidos não possui um efeito direto ou mensurável no *rating ESG* das empresas, isso pode indicar que a adoção de práticas sustentáveis e o comprometimento com as métricas ESG constituem abordagens globais, aplicáveis em diferentes partes do mundo e que transcende as fronteiras geográficas. Além disso, essa constatação também suscita a hipótese de que outros fatores como políticas regulatórias

internas ou pressões por partes interessadas, possam exercer um impacto mais acentuado nesse contexto específico.

Já a ausência de significância estatística na variável “*DivYield*” sugere que, embora o rendimento de dividendos seja fundamental no contexto financeiro, pode não ter uma ligação direta com o desempenho ESG das empresas. Isso pode refletir uma desconexão entre a distribuição de dividendos aos acionistas e as práticas de sustentabilidade adotadas pela empresa, o que implica que as empresas podem adotar abordagens distintas para obter ganhos financeiros e, ao mesmo tempo, criar impacto social e ambiental, podendo ser independentes.

De seguida, relativamente à variável “*PeerGroup*”, a falta de significância estatística, ressalta que o tipo de indústria ou setor no qual uma empresa opera não determina diretamente seu *rating ESG*, o que se traduz na necessidade de examinar mais a fundo como as empresas equilibram as suas estratégias de crescimento com compromissos ESG.

Após a análise de regressão, o foco voltou-se para uma análise mais aprofundada das variáveis mais significativas e a sua influência no desempenho ESG das empresas, sendo que os dois modelos destacaram as mesmas variáveis, que apresentaram uma forte correlação estatisticamente significativa, a “*Sus_E*”, “*Sus_S*”, “*Sus_G*”, “*R_E*”, “*R_S*”, “*R_G*” e “*Region_Europe*”, sugerindo uma consistência notável nas conclusões entre os modelos, destacando o papel crucial das práticas sustentáveis e a influência regional na determinação do *rating ESG* das empresas.

Em relação à capacidade preditiva dos modelos, os resultados desta dissertação, apontam para uma notável diferença entre os dois modelos analisados, sendo o modelo da *Sustainalytics* mais consistente e preciso em relação ao modelo do *Refinitiv*, o que pode ser considerado uma vantagem em termos de confiabilidade e aplicabilidade prática na avaliação de práticas sustentáveis e responsáveis das empresas. O segundo modelo, apresentou variações mais acentuadas nas suas previsões e resultados, o que pode indicar uma menor capacidade de capturar com precisão as influências e interações das variáveis em questão, sendo menos eficaz.

De seguida, o estudo realça a contribuição crucial da área da matemática financeira no avanço das análises ESG, enquanto otimiza as abordagens matemáticas na avaliação financeira. Uma das contribuições mais significativas deste estudo é a construção e validação de um modelo de regressão multivariado capaz de reforçar a sua importância na atribuição de *rating ESG* às empresas, assemelhando-se ao interesse do *rating*. Essa abordagem metodológica representa um avanço notável, oferecendo uma base sólida para a avaliação de riscos e oportunidades associadas a fatores E, S e G mas também uma compreensão mais detalhada das relações subjacentes que moldam as avaliações, orientando investidores e tomadores de decisão a considerar informações mais precisas nas suas estratégias futuras.

Posteriormente, embora a pesquisa tenha alcançado resultados apreciáveis, é impreterível reconhecer as *limitações* inerentes ao estudo. Estas limitações destacam áreas que podem ser alvo de melhorias em investigações futuras a fim de enriquecer ainda mais o campo do investimento sustentável e responsável.

Uma das principais limitações diz respeito à disponibilidade e qualidade dos dados utilizados, embora se tenham aplicado esforços para obter um conjunto de dados abrangentes e atualizados, algumas variáveis relevantes podem não ter sido consideradas devido a limitações de acesso ou qualidade. Uma das principais preocupações, é de que os *ratings ESG* possam ser manipulados pelas empresas para melhorar a sua imagem e reputação. Uma das melhorias futuras, pode incluir o acesso a bases de dados mais completas e atualizadas, permitindo uma análise mais abrangente e precisa das variáveis envolvidas.

Além disso, a metodologia adotada neste estudo baseou-se, principalmente, em modelos de regressão linear, que podem não capturar totalmente a complexidade das relações entre as variáveis. Abordagens mais avançadas, como modelos de *machine learning* e inteligência artificial, podem ser exploradas para aprimorar a previsão do desempenho ESG e identificar padrões não lineares ou interações mais sutis entre as variáveis.

Do mesmo modo, esta dissertação concentrou-se principalmente em métodos quantitativos, a incorporação de abordagens qualitativas, como estudos de caso ou entrevistas com especialistas do tema, pode enriquecer a compreensão das complexas interações entre ESG e desempenho corporativo.

No entanto, essas limitações não devem ser vistas como obstáculos, mas sim como um ponto de partida para futuras pesquisas, destacando a necessidade de um conjunto mais abrangente de dados e metodologias avançadas.

Nesse sentido, considera-se que a uniformização do método, aliada ao potencial das novas tecnologias como a análise de *big data* e inteligência artificial, poderia proporcionar uma avaliação mais completa e fundamentada do desempenho ESG das empresas, incorporando informações de diversas fontes, como estratégias de comunicação e alternativas.

Neste espírito, esta conclusão marca, tanto um ponto de chegada quanto um ponto de partida. À medida que a pesquisa nesta área evolui, espera-se que este estudo continue a inspirar investigações e a contribuir para a construção de um ambiente financeiro global mais consciente e alinhado com valores éticos e sustentáveis, fornecendo orientações práticas para a tomada de decisões informadas e sustentáveis.

Referências Bibliográficas

- Alareeni, B. A., & Hamdan, A. (2020). ESG impact on performance of US S&P 500-listed firms. *Corporate Governance: The International Journal of Business in Society*.
- Azhar, N. A., Pan, G., Seow, P.-S., Koh, A. e Tay, W.-Y. (2019). Abordagem de análise de texto para examinar a responsabilidade social corporativa. *Asian Journal of Accounting and Governance*, 11, 85-96.
- Bak, C., Bhattacharya, A., Edenhofer, O., & Knopf, B. (2017). Towards a comprehensive approach to climate policy, sustainable infrastructure, and finance. *Economics*, 11(1), 1-13. DOI: 10.5018/economics-ejournal.ja.2017-33.
- Berg, F., Kölbel, J. F., & Rigobon, R. (2022). Aggregate confusion: The divergence of ESG ratings. *Revista de Finanças Aplicadas*, 12(1), 1-27. Disponível em: <http://revfin.org/aggregate-confusion-the-divergence-of-esg-ratings/>. Acedido em março de 2023.
- Caixa Geral de Depósitos. (2022). Financiamento Sustentável. Disponível em: <https://www.cgd.pt/Sustentabilidade/Negocio-Responsavel/Pages/Financiamento-Sustentavel.aspx>. Acedido em julho de 2023.
- Caixa Geral de Depósitos. (n.d.). Modelo Rating ESG. Disponível em: <https://www.cgd.pt/Sustentabilidade/Negocio-Responsavel/Pages/Modelo-Rating-ESG.aspx>. Acedido em julho de 2023.
- Comissão Europeia. (2020). Comunicação Da Comissão Ao Parlamento Europeu, Ao Conselho, Ao Comité Económico E Social Europeu E Ao Comité Das Regiões - Reforçar a ambição climática da Europa para 2030. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:52020DC0562&from=pt>. Acedido em maio de 2023.
- Comissão do Mercado de Valores Mobiliários. (n.d.). Regulação em sustentabilidade. Disponível em: https://www.cmvm.pt/pt/SDI/sustentabilidade/Pages/regulacao_sustentabilidade.aspx. Acedido em maio de 2023.
- Domini, A., & Kinder, P. (1986). *Ethical Investing*. Simon and Schuster.
- Dorfleitner, G., Halbritter, G., & Nguyen, M. (2015). Measuring the level and risk of corporate responsibility - an empirical comparison of different ESG rating approaches. *Journal of Asset Management*, 16.
- Drempetic, S., Klein, C., & Zwergel, B. (2020). The influence of firm size on the ESG score: Corporate sustainability ratings under review. *Journal of Business Ethics*, 167.
- Duque-Grisales, E., & Aguilera-Caracuel, J. (2021). Environmental, social and governance (ESG) scores and financial performance of multilatinas: Moderating effects of geographic

- international diversification and financial slack. *Journal of Business Ethics*, 168(2), 315-334.
- Dyck, A., Volchkova, N., & Zingales, L. (2008). The corporate governance role of the media. *Journal of Financial Economics*, 88(1), 1-22.
- Elkington, J. (2018). *Cannibals with forks: The triple bottom line of 21st century business*. Routledge.
- Erhart, S. (2022). Take it with a pinch of salt—ESG rating of stocks and stock indices. *International Review of Financial Analysis*, 83, 102308. DOI: <https://doi.org/10.1016/j.irfa.2022.102308>
- ERM. (2019). Rate the raters 2019: Experts views on ESG ratings. Disponível em: <https://sustainability.com/our-work/reports/rate-raters-2019/>. Acedido em maio de 2023.
- European Central Bank. (n.d.). Euro foreign exchange reference rates: Chinese yuan (CNY). Disponível em: https://www.ecb.europa.eu/stats/policy_and_exchange_rates/euro_reference_exchange_rates/html/eurofxref-graph-cny.pt.html. Acedido em julho de 2023.
- Fang, M., Nie, H., & Shen, X. (2022). Can enterprise digitization improve ESG performance? *Economic Modelling*. DOI: <https://doi.org/10.1016/j.econmod>.
- Färber, M., Aue, T., & Jatowt, A. (2020). Multivariate Time Series Analysis for ESG Rating Prediction Using News Articles. *Information Retrieval Journal*, 1-29. DOI: 10.1007/s10791-020-09432-3
- Fischer, T. M., & Sawczyn, A. A. (2013). The relationship between corporate social performance and corporate financial performance and the role of innovation: Evidence from German listed firms. *Journal of management control*, 24(1), 27-52.
- Friede, G., Busch, T., & Bassen, A. (2015). ESG and financial performance: aggregated evidence from more than 2000 empirical studies. *Journal of Sustainable Finance & Investment*, 5(4), 210-233. DOI: 10.1080/20430795.2015.1118917
- FundsPeople. (2022). Cresce a convicção dos investidores portugueses sobre a atratividade dos fundos sustentáveis. FundsPeople. Disponível em: <https://fundspeople.com/pt/cresce-a-conviccao-dos-investidores-portugueses-sobre-a-atratividade-dos-fundos-sustentaveis/>. Acedido em junho de 2023.
- García, J., González-Bueno, J., Oliver, R., & Riley (2019). Selecting Socially Responsible Portfolios: A Fuzzy Multicriteria Approach. *Sustainability*, 11(9), 2496. DOI:10.3390/su11092496
- Garz, H. & Volk, C. (2018). The ESG risk ratings moving up the innovation curve white paper—Volume1. *Sustainalytics*. Disponível em: <https://connect.sustainalytics.com/hubfs/INV%20%20Reports%20and%20Brochure/Thou>

- ght%20Leadership/SustainalyticsESGRiskRatings_WhitePaperVolumeOne_October%202018.pdf. Acedido em julho de 2023.
- Global Sustainable Investment Alliance. (2020). 2020 Global Sustainable Investment Review. Disponível em: https://www.gsi-alliance.org/wp-content/uploads/2021/03/GSIR_Review2020.3.31.pdf. Acedido em abril de 2023.
- Grewal, J. S. (2019). Environmental, Social, and Governance (ESG) Performance and Financial Performance: A Meta-Analysis. *Journal of Business Ethics*, 1-28. DOI: 10.1007/s10551-019-04218-8
- Hartzmark, S. M., & Sussman, A. B. (2019). Do investors value sustainability? A natural experiment examining ranking and fund flows. *Journal of Finance*, 74(6), 2789-2837.
- Hisano, R., Sornette, D. e Mizuno, T. (2020). Previsão de conformidade ESG usando uma rede de informações heterogêneas. *Journal of Big Data*, 7(1), 1-19. DOI: 10.1186/s40537-020-00363-9
- Informa D&B. (2022). ESG Intelligence. Disponível em: <https://www.informadb.pt/pt/areas-de-negocio/esg-intelligence/>. Acedido em maio de 2023.
- Inverco. (2018). Las instituciones de inversión colectiva y los fondos de pensiones. Disponível em: <http://www.inverco.es/archivosdb/c87-ahorro-financiero-de-las-familias-iics-y-fp-2017.pdf>. Acedido em maio de 2023.
- Kiriu, T., & Nozaki, M. (2020). A Text Mining Model to Evaluate Firms' ESG Activities: An Application for Japanese Firms. *Asia-Pacific Financial Markets*, 27(4). DOI:10.1007/s10690-020-09309-1.
- Kouwenberg, R., & Phunnarungsi, W. (2013). Social media and corporate disclosures: Evidence from China. *European Financial Management*, 19(5), 921-942.
- KPMG International. (2018). The road ahead: The KPMG survey of corporate responsibility reporting 2017. KPMG International Cooperative.
- Liern, V., & Pérez-Gladish, B. (2018). Sustainability and social responsibility ratings: Measuring ESG performance. *Journal of Business Ethics*, 148(1), 179-194.
- Lu, Y. C., Wang, L., & Zhang, Y. (2022). Does digital financial inclusion matter for firms' ESG disclosure? Evidence from China. *Frontiers in Environmental Science*, 10, 1029975. DOI: <https://doi.org/10.3389/fenvs.2022.1029975>
- Lydenberg, S., & Domini, A. (1993). Investing for Good: Making Money While Being Socially Responsible. Harper Collins.
- Maiti, M. (2020). Is ESG the succeeding risk factor? *Journal of Sustainable Finance & Investment*, 1-15. DOI: 10.1080/20430795.2020.1775265
- Morningstar. (2023). A New Paradigm for Understanding Risk, de Thomas Kuh, Feb 21, 2023. Disponível em: <https://www-morningstar-com.translate.goog/articles/1139118/a-new->

- paradigm-for-understanding-risk?_x_tr_sl=en&_x_tr_tl=pt&_x_tr_hl=pt-PT&_x_tr_pto=sc&_x_tr_hist=true. Acedido em abril de 2023.
- Morningstar. (2022). Global Sustainable Fund Flows: Q1 2022 in Review. Disponível em: <https://www.morningstar.com/articles/1091550/us-sustainable-fund-flows-slid-in-first-quarter-2022>. Acedido em abril de 2023.
- Nollet, J., Filis, G., & Mitrokostas, E. (2016). Corporate social responsibility and financial performance: A non-linear and disaggregated approach. *Economic Modelling*, 52, 400-407.
- Orecchio, A. (2022). ESG rating and rankings: Meaning and definition. FinScience. Disponível em: <https://finscience.com/en/blog/esg/esg-rating-and-rankings-meaning-and-definition/>. Acedido em maio de 2023.
- Pacto Global das Nações Unidas. (2004). Who Cares Wins: Connecting Financial Markets to a Changing World: Report of the World Economic Forum Investors Summit on Climate Risk. Disponível em: <https://www.unglobalcompact.org/library/56>. Acedido em abril maio de 2023.
- Quivy, R., & Luc Campenhoudt, L. (2008). Manual de investigação em ciências sociais. Lisboa. Gradiva.
- Refinitiv. (2022). Refinitiv ESG company scores. Disponível em: <https://www.refinitiv.com/en/sustainable-finance/esg-scores>. Acedido em abril de 2023.
- Sachs, J. D. (2015). The age of sustainable development. Columbia University Press.
- Sætra, H. (2021). A Framework for Evaluating and Disclosing the ESG Related Impacts of AI with the SDGs. *Sustainability*, 13(15), 8503.
- Schmutz, B., Tehrani, M., Fulton, L., & Rathgeber, A. (2020). Dow Jones sustainability indices, do they make a difference? The US and the European Union companies. *Sustainability*, 12(17), 6785.
- Sievänen, R., Sumelius, J., Islam, K., & Sell, M. (2013). From struggle in responsible investment to potential to improve global environmental governance through UN PRI. *International Environmental Agreements: Politics, Law and Economics*, 13(2), 197-217.
- Singh, A., Castro, P., & Archer, D. (2012). An overview of sustainability assessment methodologies. *Ecological Indicators*, 15(3), 281-299.
- Sorrosal-Forradas, M.-T., Barberà-Mariné, M.-G., Fabregat-Aibar, L., & Li, X. (2023). A new rating of sustainability based on the Morningstar Sustainability Rating. *Journal of Cleaner Production*, 306, 127164. DOI: 10.1016/j.jclepro.2021.127164.
- Swinburne, J., Pillay, S., & Orlov, A. (2019). The Changing Landscape of Investor Focus: Beyond the Financials. *Journal of Sustainable Finance & Investment*, 9(2), 112-125. DOI: <https://doi.org/10.1080/20430795.2019.1580553>

- UNRIC. (s.d.). Objetivos de Desenvolvimento Sustentável. United Nations Regional Information Centre for Western Europe. Disponível em: de <https://unric.org/pt/Objetivos-de-Desenvolvimento-Sustentavel/>. Acedido em fevereiro de 2023.
- United Nations. (2015). Transforming our world: The 2030 agenda for sustainable development. United Nations General Assembly.
- Velte, P. (2017). Does ESG performance have an impact on financial performance? Evidence from Germany. *Journal of Global Responsibility*.
- Waddock, S. A., & Graves, S. B. (1997). The corporate social performance–financial performance link. *Strategic management journal*, 18(4), 303-319.
- Wang, W., Sun, Z., & Wang, W. (2023). The impact of environmental uncertainty on ESG performance: Emotional vs. rational. *Journal of Cleaner Production*, 397, 136528. <https://doi.org/10.1016/j.jclepro.2023.136528>
- Xie, J., Nozawa, W., Yagi, M., Fujii, H., & Managi, S. (2019). Do environmental, social, and governance activities improve corporate financial performance?. *Business Strategy and the Environment*, 28(2), 286-300.
- Yi Yu, E. P., & Luu, B. V. (2021). International variations in ESG disclosure – do cross-listed companies care more? *International Review of Financial Analysis*, 75, Article 101731.

Anexos

Anexo A- Escala de Rating

Risco	Fitch Ratings	Moody's	S&P Global Ratings	Categoria
	AAA	Aaa	AAA	Grau de investimento
	AA+	Aa1	AA+	
	AA	Aa2	AA	
	AA-	Aa3	AA-	
	A+	A1	A+	
	A	A2	A	
	A-	A3	A-	
	BBB+	Baa1	BBB+	
	BBB	Baa2	BBB	
	BBB-	Baa3	BBB-	Grau especulativo
	BB+	Ba1	BB+	
	BB	Ba2	BB	
	BB-	Ba3	BB-	
	B+	B1	B+	
	B	B2	B	
	B-	B3	B-	Vulnerabilidade a não pagamento
	CCC	Caa1	CCC	
	CC	Caa2	CC	
	C	Caa3	C	
	D	Ca C	D	

Fonte: XP Investment

Anexo B- Evolução do ESG



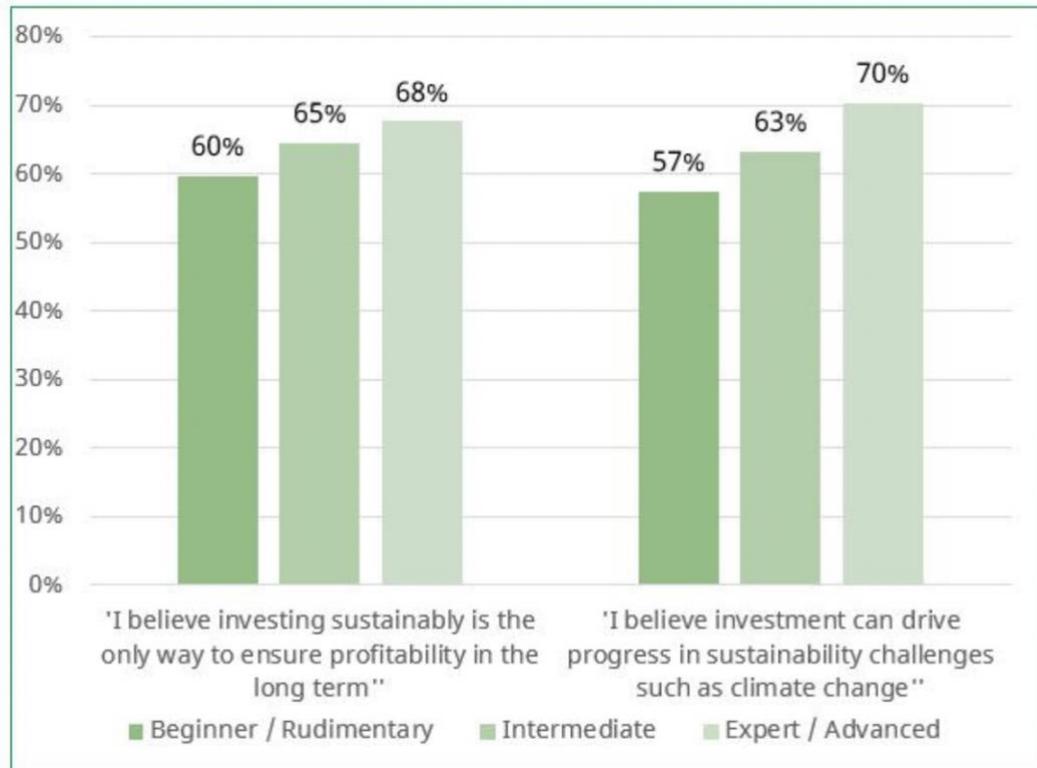
Fonte: Google Trends

Anexo C- Objetivos de Desenvolvimento Sustentável (ODS)

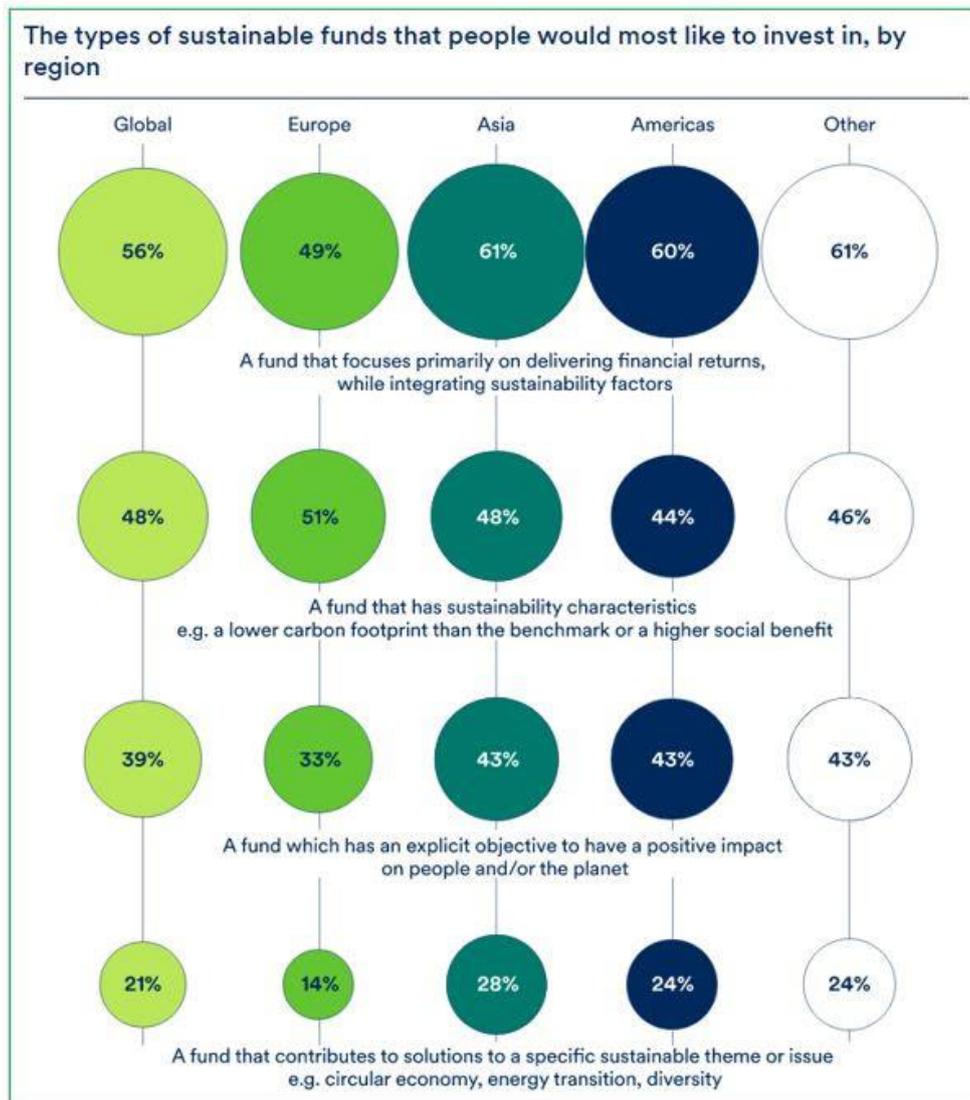


Fonte: Nações Unidas

Anexo D- Estudo *Schroders Global Investor Study 2022*



Fonte: FundsPeople



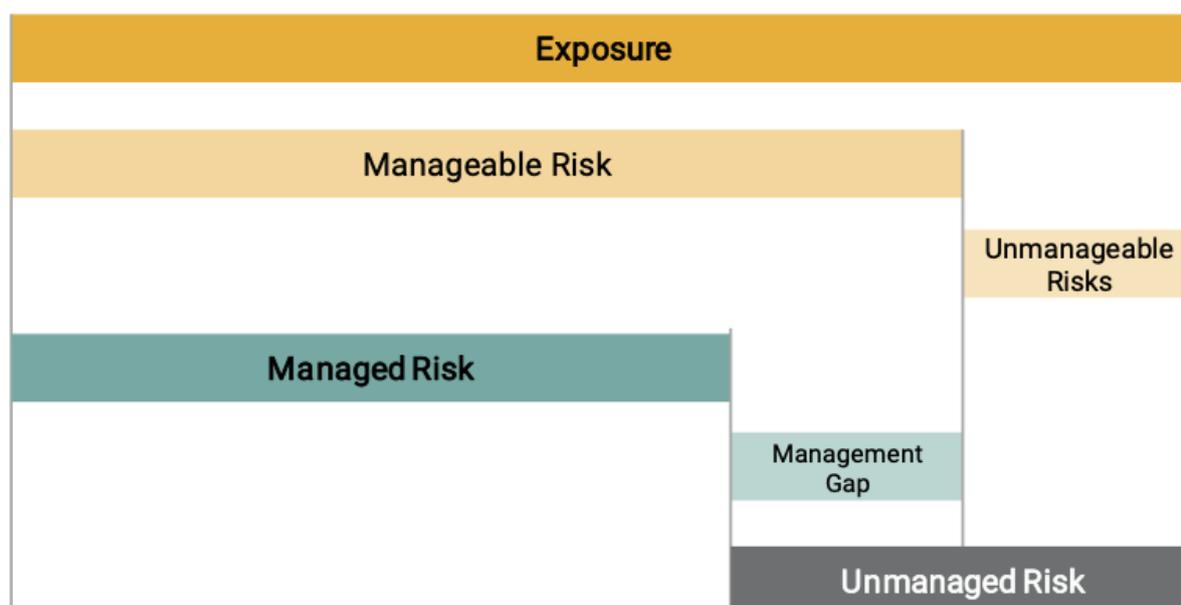
Fonte: FundsPeople

Anexo E- Interpretação dos scores ESG da Refinitiv

Score	Descrição
0-25	Indica um desempenho ESG relativo fraco e grau insuficiente de transparência na divulgação pública de dados ESG.
25-50	Indica um desempenho ESG relativo satisfatório e grau moderado de transparência na divulgação pública de dados ESG materiais.
50-75	Indica um desempenho ESG relativo bom e grau acima da média de transparência na divulgação pública de dados ESG materiais.
75-100	Indica um desempenho ESG relativo excelente e alto grau de transparência na divulgação pública de dados ESG materiais.

Fonte: Adaptado pelo autor da Refinitiv

Anexo F- Decomposição do risco da *Sustainalytics*



Fonte: Sustainalytics

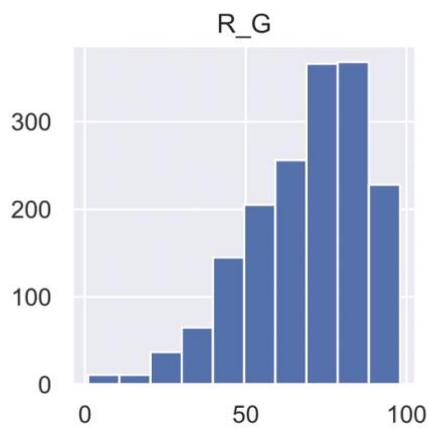
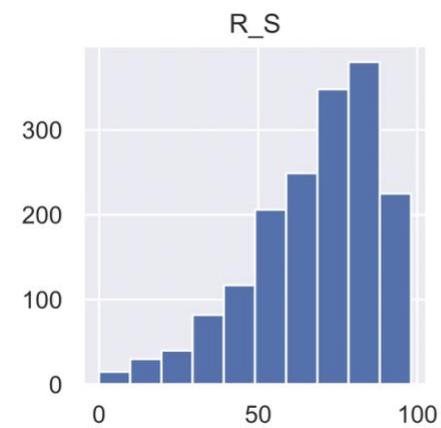
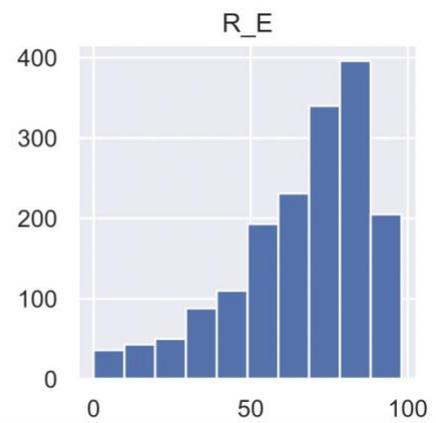
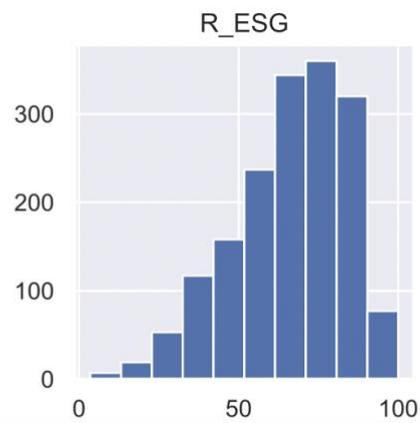
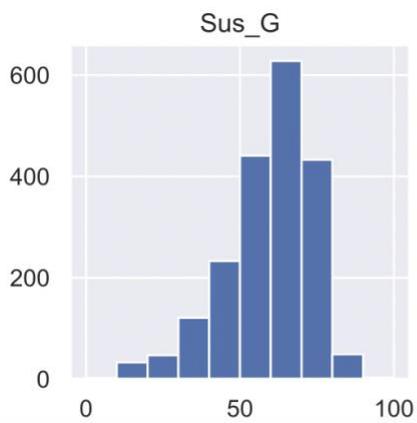
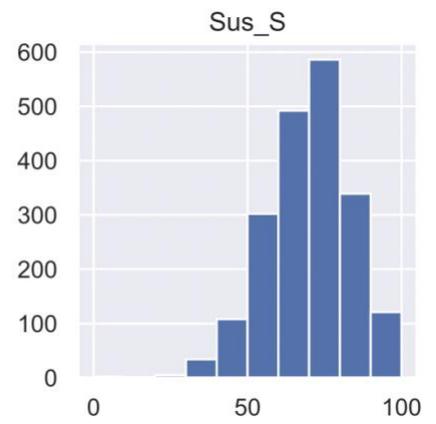
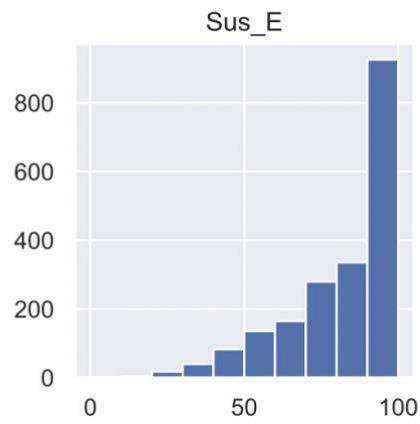
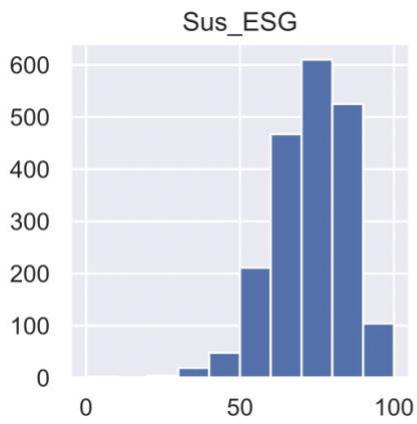
Anexo G-Descrição de variáveis

Nome das variáveis	Descrição	Tipo
Company	Representa o nome ou identificação única da empresa.	Categórica
Sus_ESG	Refere-se ao rating ESG da empresa, calculado com a metodologia da <i>Sustainalytics</i> .	Numérica
Sus_E, Sus_S, Sus_G	São os ratings E, S e G, que representam as dimensões Ambiental (E), Social (S) e de <i>Governance</i> (G) no contexto da análise ESG. Cada um deles é um componente do <i>rating ESG</i> total.	Numérica
R_ESG	Representa o rating ESG da empresa, calculado com a metodologia da <i>Refinitiv</i> .	Numérica
R_E, R_S, R_G	São os ratings E, S e G, que representam as dimensões Ambiental (E), Social (S) e de <i>Governance</i> (G) no contexto da análise ESG. Cada um deles é um componente do <i>rating ESG</i> total.	Numérica
Year, Month	Indicam o ano e o mês associados aos dados. Como o foco do estudo é analisar a relação entre os ratings ESG e vários fatores, as informações de ano e mês não são consideradas	Numérica/ Categórica

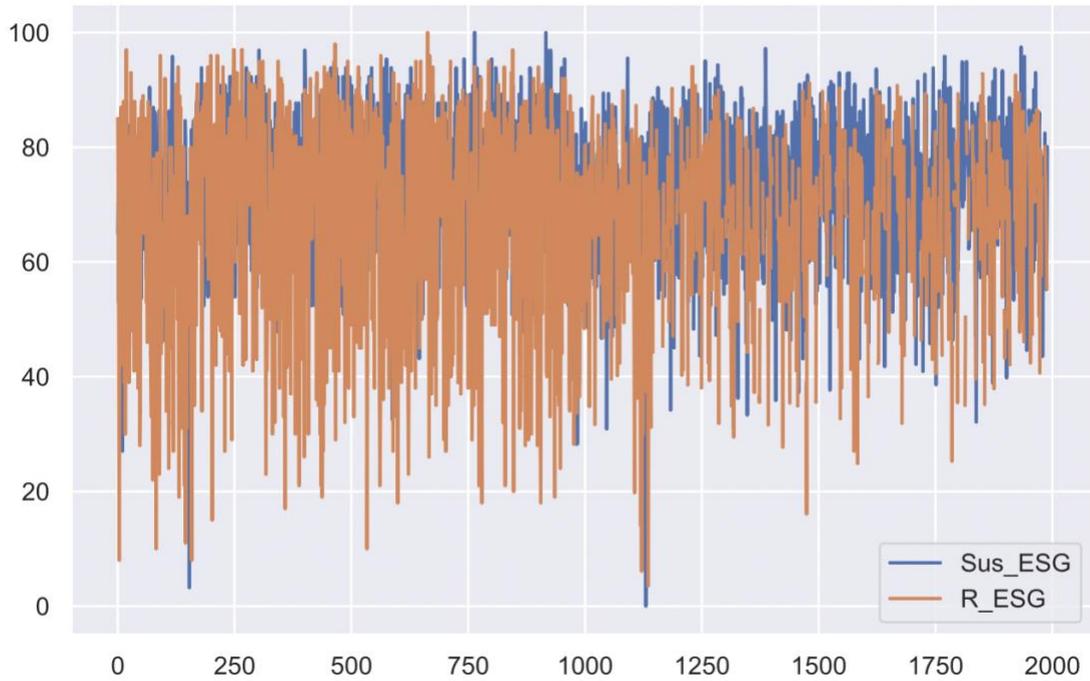
	diretamente relevantes para a análise. Portanto, essas variáveis foram deixadas de fora para manter o foco nas dimensões de interesse.	
PeerGroup	Refere-se ao grupo de atividade ao qual a empresa pertence, utilizado para comparar o seu desempenho.	Categórica
Index	Indica o índice bolsista ao qual a empresa está vinculada.	Categórica
Contry	Representa o país no qual a empresa está sediada.	Categórica
Region	Indica a região geográfica à qual o país pertence.	Categórica
MarkCap	É o peso de mercado da empresa.	Numérica
DivYield	Refere-se ao rendimento de dividendos da empresa.	Numérica
BookValue	Representa o valor contabilístico da empresa.	Numérica
200D, PB, LTMPE	São possíveis métricas financeiras ou indicadores utilizados na análise. Não foram incluídas no modelo por duas razões principais. Primeiro, para evitar a multicolinearidade, que ocorre quando as variáveis independentes estão altamente correlacionadas. Segundo, para manter o modelo mais simples e interpretável, focando nas variáveis que estão mais relacionadas às dimensões ESG.	Numérica
CountryCds	Representa o <i>Credit Default Swaps</i> (CDS) associados ao país da empresa. Embora essa variável possa fornecer informações sobre a saúde financeira do país, ela não é o foco central da análise ESG e financeira. Portanto, foi excluída para manter a análise mais direcionada aos aspetos da investigação.	Numérica
Sus_E_n_r Sus_S_n_r, Sus_G_n_r	Podem ser valores normalizados dos <i>ratings</i> E, S e G da <i>Sustainalytics</i> . Embora esses valores normalizados dos <i>ratings</i> ESG da <i>Sustainalytics</i> possam ser informativos, optou-se por não os incluir no modelo para manter a análise mais focada nos <i>ratings</i> não normalizados e em outras variáveis que possam ter uma influência mais direta no desempenho ESG.	Numérica
B_E, B_S, B_G	São pesos atribuídos às dimensões E, S e G na metodologia da <i>Bloomberg</i> . Como esses pesos atribuídos às dimensões E, S e G na metodologia da <i>Bloomberg</i> são características específicas da metodologia utilizada e apresentavam mais de metade das observações em valores omissos, optou-se por não incluir essas variáveis para evitar a introdução de complexidade adicional.	Numérica

Fonte: Autor

Anexo H- Histogramas



Anexo I- Diferenças e semelhanças do rating ESG nas duas agências



Anexo J- Modelo I: Fatores E, S, G e o Rating ESG

OLS Regression Results

Dep. Variable:	Sus_ESG	R-squared:	0.971			
Model:	OLS	Adj. R-squared:	0.971			
Method:	Least Squares	F-statistic:	1.754e+04			
Date:	Thu, 06 Jul 2023	Prob (F-statistic):	0.00			
Time:	16:36:20	Log-Likelihood:	-3375.9			
No. Observations:	1579	AIC:	6760.			
Df Residuals:	1575	BIC:	6781.			
Df Model:	3					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
const	-13.2276	0.380	-34.836	0.000	-13.972	-12.483
Sus_E	0.4694	0.003	158.004	0.000	0.464	0.475
Sus_S	0.4597	0.004	103.410	0.000	0.451	0.468
Sus_G	0.2683	0.004	65.879	0.000	0.260	0.276
Omnibus:	3648.930	Durbin-Watson:	1.983			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	19276778.147			
Skew:	-21.641	Prob(JB):	0.00			
Kurtosis:	542.559	Cond. No.	904.			

OLS Regression Results

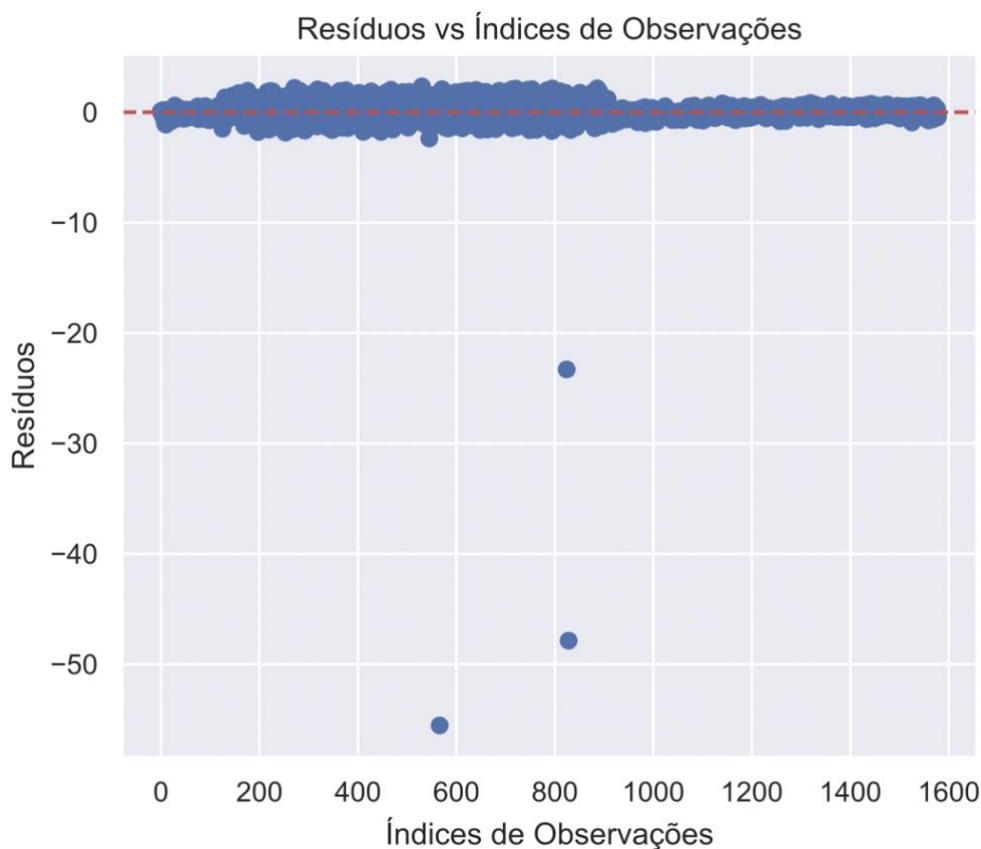
Dep. Variable:	R_ESG	R-squared:	0.301
Model:	OLS	Adj. R-squared:	0.300
Method:	Least Squares	F-statistic:	226.0
Date:	Wed, 12 Jul 2023	Prob (F-statistic):	6.14e-122
Time:	18:07:15	Log-Likelihood:	-6478.6
No. Observations:	1579	AIC:	1.297e+04
Df Residuals:	1575	BIC:	1.299e+04
Df Model:	3		
Covariance Type:	nonrobust		

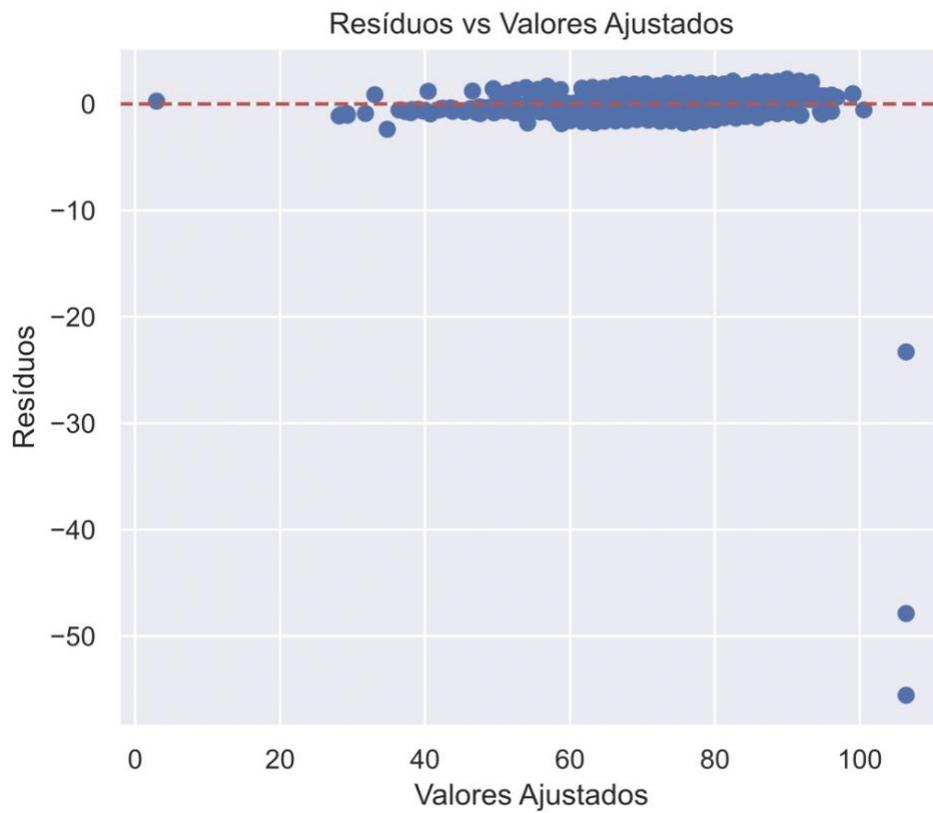
	coef	std err	t	P> t	[0.025	0.975]
const	25.7688	1.642	15.693	0.000	22.548	28.990
R_E	0.1550	0.031	4.925	0.000	0.093	0.217
R_S	0.1653	0.035	4.736	0.000	0.097	0.234
R_G	0.2782	0.023	12.282	0.000	0.234	0.323

Omnibus:	144.103	Durbin-Watson:	1.808
Prob(Omnibus):	0.000	Jarque-Bera (JB):	214.107
Skew:	-0.691	Prob(JB):	3.22e-47
Kurtosis:	4.160	Cond. No.	542.

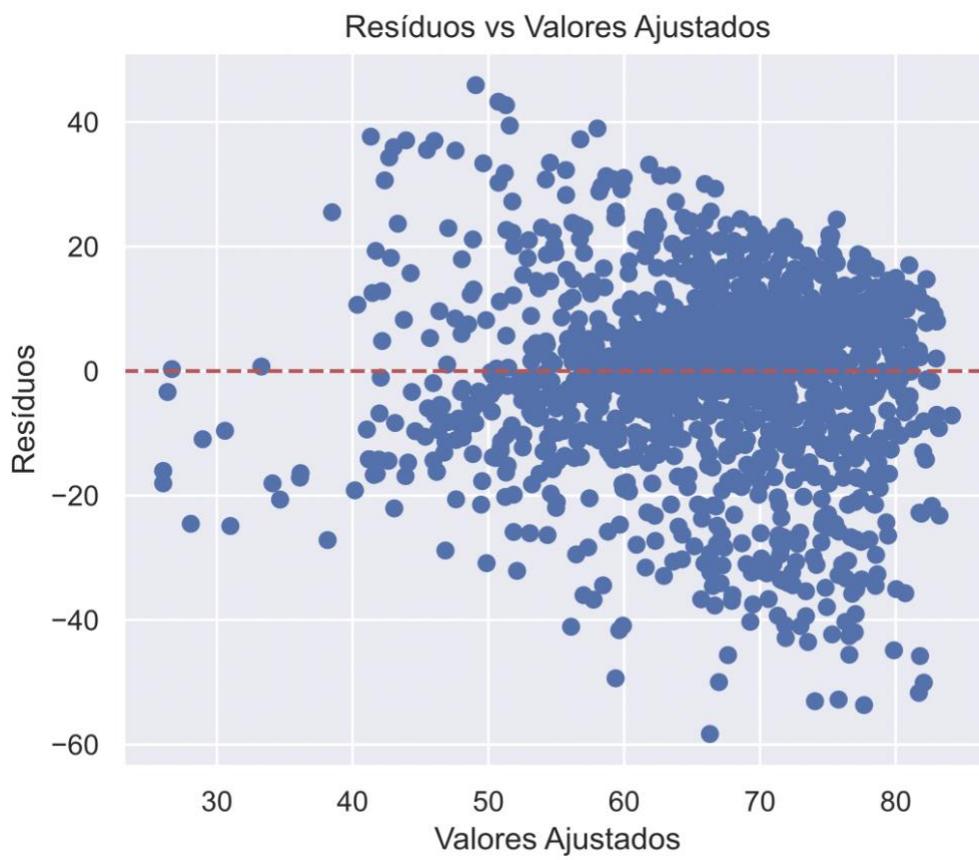
Anexo K- Resíduos do Modelo I

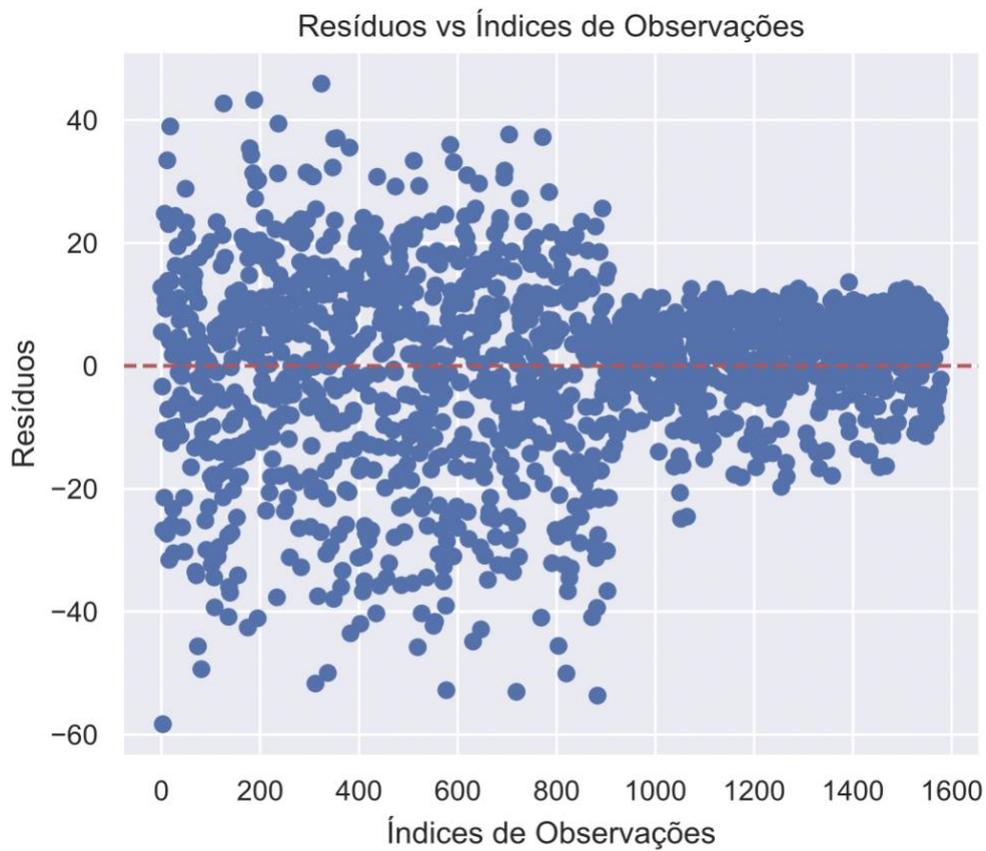
Sustainalytics:





Refinitiv:





Anexo L- Pressuposto dos resíduos

Testes/Modelos	Teste de Shapiro-Wilk (Normalidade dos Resíduos)	Teste de Breusch-Pagan (Homoscedasticidade)	Teste de Durbin-Watson (Autocorrelação)
Sustainalytics	Estatística de Teste: 0,1951	Estatística de Teste: 538,953	Estatística de Teste: 1,9807
	Valor-p: 0,0	Valor-p: 7,6815e-107	-
Refinitiv	Estatística de Teste: 0,9610	Estatística de Teste: 12,730	Estatística de Teste: 1,8079
	Valor-p: 3,5999e-20	Valor-p: 0,4688	-

Fonte: Autor

Anexo M- Relação entre variáveis categóricas

Métrica	Valor
Valor Qui-quadrado	5970,0
Valor-p	0,0
Coeficiente V de Cramer	1,0

Fonte: Autor

Anexo N- Modelo II: Fatores geográficos, financeiros, E, S, G e o Rating ESG

OLS Regression Results

```

=====
Dep. Variable:          Sus_ESG      R-squared:                0.981
Model:                  OLS          Adj. R-squared:           0.981
Method:                 Least Squares  F-statistic:              8918.
Date:                   Mon, 17 Jul 2023  Prob (F-statistic):       0.00
Time:                   15:57:02      Log-Likelihood:           -3047.6
No. Observations:      1579          AIC:                      6115.
Df Residuals:          1569          BIC:                      6169.
Df Model:               9
Covariance Type:       nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
const	-13.7058	0.322	-42.509	0.000	-14.338	-13.073
Sus_E	0.4718	0.002	192.135	0.000	0.467	0.477
Sus_S	0.4641	0.004	125.584	0.000	0.457	0.471
Sus_G	0.2719	0.004	76.110	0.000	0.265	0.279
MarkCap	5.116e-14	3.09e-14	1.656	0.098	-9.43e-15	1.12e-13
BookValue	-0.0002	5.54e-06	-28.179	0.000	-0.000	-0.000
Region_Australia	-0.1002	0.210	-0.477	0.633	-0.512	0.312
Region_Europe	-0.2765	0.147	-1.880	0.060	-0.565	0.012
Region_North America	-0.0920	0.138	-0.666	0.506	-0.363	0.179
DivYield	-0.8543	1.622	-0.527	0.598	-4.035	2.327

```

=====
Omnibus:                3934.377      Durbin-Watson:            1.982
Prob(Omnibus):           0.000        Jarque-Bera (JB):         45205723.646
Skew:                    -25.714      Prob(JB):                 0.00
Kurtosis:                830.320      Cond. No.:                6.54e+13
=====

```

OLS Regression Results

=====						
Dep. Variable:	R_ESG	R-squared:	0.316			
Model:	OLS	Adj. R-squared:	0.312			
Method:	Least Squares	F-statistic:	80.59			
Date:	Thu, 06 Jul 2023	Prob (F-statistic):	7.18e-123			
Time:	16:36:21	Log-Likelihood:	-6461.3			
No. Observations:	1579	AIC:	1.294e+04			
Df Residuals:	1569	BIC:	1.300e+04			
Df Model:	9					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

const	25.5773	1.823	14.029	0.000	22.001	29.153
R_E	0.1653	0.032	5.200	0.000	0.103	0.228
R_S	0.1395	0.035	3.981	0.000	0.071	0.208
R_G	0.2518	0.023	10.921	0.000	0.207	0.297
MarkCap	3.598e-13	2.69e-13	1.338	0.181	-1.68e-13	8.87e-13
BookValue	-0.0001	4.8e-05	-2.396	0.017	-0.000	-2.09e-05
Region_Australia	4.4381	1.834	2.420	0.016	0.841	8.035
Region_Europe	5.8178	1.269	4.583	0.000	3.328	8.308
Region_North America	1.8336	1.162	1.579	0.115	-0.445	4.112
DivYield	11.9389	13.970	0.855	0.393	-15.463	39.341
=====						
Omnibus:	142.042	Durbin-Watson:	1.826			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	207.968			
Skew:	-0.689	Prob(JB):	6.92e-46			
Kurtosis:	4.123	Cond. No.	6.49e+13			
=====						

Anexo O: Modelo III: Fatores financeiros, E, S, G e o Rating ESG

OLS Regression Results

=====						
Dep. Variable:	Sus_ESG	R-squared:	0.981			
Model:	OLS	Adj. R-squared:	0.981			
Method:	Least Squares	F-statistic:	1.597e+04			
Date:	Thu, 06 Jul 2023	Prob (F-statistic):	0.00			
Time:	16:36:21	Log-Likelihood:	-3053.3			
No. Observations:	1579	AIC:	6119.			
Df Residuals:	1573	BIC:	6151.			
Df Model:	5					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

const	-13.5791	0.314	-43.229	0.000	-14.195	-12.963
Sus_E	0.4712	0.002	193.551	0.000	0.466	0.476
Sus_S	0.4637	0.004	126.931	0.000	0.457	0.471
Sus_G	0.2694	0.003	80.564	0.000	0.263	0.276
BookValue	-0.0002	5.49e-06	-28.078	0.000	-0.000	-0.000
DivYield	-0.9636	1.621	-0.595	0.552	-4.143	2.215
=====						
Omnibus:	3920.169	Durbin-Watson:	1.975			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	44107625.003			
Skew:	-25.485	Prob(JB):	0.00			
Kurtosis:	820.200	Cond. No.	2.98e+05			
=====						

OLS Regression Results

=====						
Dep. Variable:	R_ESG	R-squared:	0.304			
Model:	OLS	Adj. R-squared:	0.302			
Method:	Least Squares	F-statistic:	137.3			
Date:	Wed, 12 Jul 2023	Prob (F-statistic):	5.52e-121			
Time:	18:00:51	Log-Likelihood:	-6475.4			
No. Observations:	1579	AIC:	1.296e+04			
Df Residuals:	1573	BIC:	1.299e+04			
Df Model:	5					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

const	25.7523	1.663	15.481	0.000	22.489	29.015
R_E	0.1554	0.032	4.933	0.000	0.094	0.217
R_S	0.1652	0.035	4.738	0.000	0.097	0.234
R_G	0.2742	0.023	12.090	0.000	0.230	0.319
BookValue	-0.0001	4.8e-05	-2.389	0.017	-0.000	-2.05e-05
DivYield	14.5548	14.044	1.036	0.300	-12.993	42.102
=====						
Omnibus:	145.925	Durbin-Watson:	1.814			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	218.647			
Skew:	-0.694	Prob(JB):	3.32e-48			
Kurtosis:	4.181	Cond. No.	2.95e+05			
=====						

Anexo P- Modelo IV: Fatores E, S, G, financeiros, geográficos, setoriais e o Rating ESG

OLS Regression Results

=====						
Dep. Variable:	Sus_ESG	R-squared:	0.981			
Model:	OLS	Adj. R-squared:	0.981			
Method:	Least Squares	F-statistic:	6676.			
Date:	Wed, 19 Jul 2023	Prob (F-statistic):	0.00			
Time:	01:07:35	Log-Likelihood:	-3047.5			
No. Observations:	1579	AIC:	6121.			
Df Residuals:	1566	BIC:	6191.			
Df Model:	12					
Covariance Type:	nonrobust					
=====						
	coef	std err	t	P> t	[0.025	0.975]

const	-13.6735	0.344	-39.800	0.000	-14.347	-13.000
Sus_E	0.4717	0.003	168.785	0.000	0.466	0.477
Sus_S	0.4638	0.004	118.129	0.000	0.456	0.471
Sus_G	0.2719	0.004	74.032	0.000	0.265	0.279
MarkCap	5.165e-14	3.1e-14	1.665	0.096	-9.18e-15	1.12e-13
BookValue	-0.0002	5.55e-06	-28.151	0.000	-0.000	-0.000
Region_Australia	-0.1004	0.211	-0.475	0.635	-0.514	0.314
Region_Europe	-0.2732	0.148	-1.840	0.066	-0.564	0.018
Region_North America	-0.0918	0.140	-0.658	0.511	-0.366	0.182
DivYield	-0.8428	1.635	-0.516	0.606	-4.049	2.364
PeerGroup_Pharmaceuticals	-0.0674	0.224	-0.301	0.763	-0.506	0.371
PeerGroup_Food Products	-0.0279	0.203	-0.138	0.891	-0.425	0.370
PeerGroup_Oil & Gas Producers	-0.0363	0.267	-0.136	0.892	-0.560	0.488
=====						
Omnibus:	3934.197	Durbin-Watson:	1.982			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	45192603.146			
Skew:	-25.711	Prob(JB):	0.00			
Kurtosis:	830.200	Cond. No.	6.59e+13			
=====						

OLS Regression Results

```

=====
Dep. Variable:          R_ESG      R-squared:          0.317
Model:                 OLS        Adj. R-squared:     0.311
Method:                Least Squares  F-statistic:        60.44
Date:                  Wed, 12 Jul 2023  Prob (F-statistic): 3.06e-120
Time:                  18:07:27     Log-Likelihood:     -6460.8
No. Observations:     1579        AIC:                1.295e+04
Df Residuals:         1566        BIC:                1.302e+04
Df Model:              12
Covariance Type:      nonrobust
=====

```

	coef	std err	t	P> t	[0.025	0.975]
const	25.5830	1.825	14.020	0.000	22.004	29.162
R_E	0.1662	0.032	5.218	0.000	0.104	0.229
R_S	0.1380	0.035	3.922	0.000	0.069	0.207
R_G	0.2502	0.023	10.823	0.000	0.205	0.296
MarkCap	3.573e-13	2.7e-13	1.325	0.185	-1.72e-13	8.86e-13
BookValue	-0.0001	4.81e-05	-2.382	0.017	-0.000	-2.02e-05
Region_Australia	4.5095	1.839	2.452	0.014	0.903	8.116
Region_Europe	5.8317	1.272	4.585	0.000	3.337	8.327
Region_North America	1.8517	1.165	1.590	0.112	-0.433	4.136
DivYield	11.5922	14.178	0.818	0.414	-16.217	39.402
PeerGroup_Pharmaceuticals	1.0022	1.824	0.549	0.583	-2.576	4.581
PeerGroup_Food Products	1.1543	1.723	0.670	0.503	-2.226	4.535
PeerGroup_Oil & Gas Producers	1.0374	2.084	0.498	0.619	-3.049	5.124

```

=====
Omnibus:              142.109    Durbin-Watson:      1.831
Prob(Omnibus):        0.000    Jarque-Bera (JB):   207.766
Skew:                 -0.690    Prob(JB):           7.66e-46
Kurtosis:             4.119    Cond. No.           6.58e+13
=====

```

Anexo Q- Testes estatísticos

Teste t de Student - Variáveis: Sus_E e Sus_ESG
 Estatística t: 14.605099595244972
 Valor p: 8.4787858064348585e-47

Teste ANOVA - Variáveis: Sus_E e Sus_ESG
 Estatística F: 213.30893418702547
 Valor p: 8.4787858064327e-47

Teste t de Student - Variáveis: Sus_E e R_ESG
 Estatística t: 21.626765263475978
 Valor p: 7.739951011805905e-97

Teste ANOVA - Variáveis: Sus_E e R_ESG
 Estatística F: 467.71697576149126
 Valor p: 7.739951011806655e-97

Teste t de Student - Variáveis: Sus_S e Sus_ESG
 Estatística t: -6.396833741763782
 Valor p: 1.8208296165476275e-10

Teste ANOVA - Variáveis: Sus_S e Sus_ESG
 Estatística F: 40.9194819197676
 Valor p: 1.820829616547606e-10

Teste t de Student - Variáveis: Sus_S e R_ESG

Estatística t: 5.091119490801803
Valor p: 3.768184257703989e-07

Teste ANOVA - Variáveis: Sus_S e R_ESG
Estatística F: 25.9194976696219
Valor p: 3.7681842577052217e-07

Teste t de Student - Variáveis: Sus_G e Sus_ESG
Estatística t: -27.170361021368926
Valor p: 2.9243062794331843e-146

Teste ANOVA - Variáveis: Sus_G e Sus_ESG
Estatística F: 738.2285180315233
Valor p: 2.924306279434569e-146

Teste t de Student - Variáveis: Sus_G e R_ESG
Estatística t: -12.595148188466048
Valor p: 1.5838411907165613e-35

Teste ANOVA - Variáveis: Sus_G e R_ESG
Estatística F: 158.6377578894196
Valor p: 1.5838411907168778e-35

Teste t de Student - Variáveis: DivYield e Sus_ESG
Estatística t: -239.1414735322484
Valor p: 0.0

Teste ANOVA - Variáveis: DivYield e Sus_ESG
Estatística F: 57188.64436317507
Valor p: 0.0

Teste t de Student - Variáveis: DivYield e R_ESG
Estatística t: -151.4775856061335
Valor p: 0.0

Teste ANOVA - Variáveis: DivYield e R_ESG
Estatística F: 22945.4589410635
Valor p: 0.0

Teste t de Student - Variáveis: BookValue e Sus_ESG
Estatística t: 3.44553112576151
Valor p: 0.0005773770352894628

Teste ANOVA - Variáveis: BookValue e Sus_ESG
Estatística F: 11.871684738591384
Valor p: 0.0005773770352896425

Teste t de Student - Variáveis: BookValue e R_ESG
Estatística t: 3.474843659912404
Valor p: 0.0005180375329059963

Teste ANOVA - Variáveis: BookValue e R_ESG
Estatística F: 12.074538460833432
Valor p: 0.000518037532906006

Teste t de Student - Variáveis: MarkCap e Sus_ESG
Estatística t: 12.725406935440699
Valor p: 3.2653465866549356e-36

Teste ANOVA - Variáveis: MarkCap e Sus_ESG
Estatística F: 161.9359816725622
Valor p: 3.265346586654875e-36

Teste t de Student - Variáveis: MarkCap e R_ESG
Estatística t: 12.72540693558051
Valor p: 3.2653465810952075e-36

Teste ANOVA - Variáveis: MarkCap e R_ESG
Estatística F: 161.93598167612055
Valor p: 3.265346581095334e-36

Teste t de Student - Variáveis: Region_Australia e Sus_ESG
Estatística t: -238.97626153291324
Valor p: 0.0

Teste ANOVA - Variáveis: Region_Australia e Sus_ESG
Estatística F: 57109.65357624701
Valor p: 0.0

Teste t de Student - Variáveis: Region_Australia e R_ESG
Estatística t: -151.3820614773751
Valor p: 0.0

Teste ANOVA - Variáveis: Region_Australia e R_ESG
Estatística F: 22916.528537139748
Valor p: 0.0

Teste t de Student - Variáveis: Region_Europe e Sus_ESG
Estatística t: -238.16501245973748
Valor p: 0.0

Teste ANOVA - Variáveis: Region_Europe e Sus_ESG
Estatística F: 56722.57315994651
Valor p: 0.0

Teste t de Student - Variáveis: Region_Europe e R_ESG
Estatística t: -150.86920615508598
Valor p: 0.0

Teste ANOVA - Variáveis: Region_Europe e R_ESG
Estatística F: 22761.5173658658
Valor p: 0.0

Teste t de Student - Variáveis: Region_North America e Sus_ESG
Estatística t: -237.48226592560366
Valor p: 0.0

Teste ANOVA - Variáveis: Region_North America e Sus_ESG
Estatística F: 56397.82662915875
Valor p: 0.0

Teste t de Student - Variáveis: Region_North America e R_ESG
Estatística t: -150.4159083606815
Valor p: 0.0

Teste ANOVA - Variáveis: Region_North America e R_ESG

Estatística F: 22624.945487968882
Valor p: 0.0

Teste t de Student - Variáveis: PeerGroup_Chemicals e Sus_ESG
Estatística t: -239.05298889258242
Valor p: 0.0

Teste ANOVA - Variáveis: PeerGroup_Chemicals e Sus_ESG
Estatística F: 57146.33149847732
Valor p: 0.0

Teste t de Student - Variáveis: PeerGroup_Chemicals e R_ESG
Estatística t: -151.42935441944962
Valor p: 0.0

Teste ANOVA - Variáveis: PeerGroup_Chemicals e R_ESG
Estatística F: 22930.84937989133
Valor p: 0.0

Teste t de Student - Variáveis: PeerGroup_Oil & Gas Producers e Sus_ESG
Estatística t: -239.08900207151163
Valor p: 0.0

Teste ANOVA - Variáveis: PeerGroup_Oil & Gas Producers e Sus_ESG
Estatística F: 57163.55091155073
Valor p: 0.0

Teste t de Student - Variáveis: PeerGroup_Oil & Gas Producers e R_ESG
Estatística t: -151.45148637701462
Valor p: 0.0

Teste ANOVA - Variáveis: PeerGroup_Oil & Gas Producers e R_ESG
Estatística F: 22937.552725806974
Valor p: 0.0

Teste t de Student - Variáveis: PeerGroup_Pharmaceuticals e Sus_ESG
Estatística t: -239.04785491439284
Valor p: 0.0

Teste ANOVA - Variáveis: PeerGroup_Pharmaceuticals e Sus_ESG
Estatística F: 57143.87693917225
Valor p: 0.0

Teste t de Student - Variáveis: PeerGroup_Pharmaceuticals e R_ESG
Estatística t: -151.42619592883503
Valor p: 0.0

Teste ANOVA - Variáveis: PeerGroup_Pharmaceuticals e R_ESG
Estatística F: 22929.892813477894
Valor p: 0.0

Teste t de Student - Variáveis: PeerGroup_Precious Metals e Sus_ESG
Estatística t: -239.18481485391314
Valor p: 0.0

Teste ANOVA - Variáveis: PeerGroup_Precious Metals e Sus_ESG
Estatística F: 57209.37565670117
Valor p: 0.0

Teste t de Student - Variáveis: PeerGroup_Precious Metals e R_ESG
 Estatística t: -151.510167840353
 Valor p: 0.0

Teste ANOVA - Variáveis: PeerGroup_Precious Metals e R_ESG
 Estatística F: 22955.33095901194
 Valor p: 0.0

Teste t de Student - Variáveis: PeerGroup_Refiners & Pipelines e Sus_ESG
 Estatística t: -239.16402044434508
 Valor p: 0.0

Teste ANOVA - Variáveis: PeerGroup_Refiners & Pipelines e Sus_ESG
 Estatística F: 57199.428675103154
 Valor p: 0.0

Teste t de Student - Variáveis: PeerGroup_Refiners & Pipelines e R_ESG
 Estatística t: -151.49745660132544
 Valor p: 0.0

Teste ANOVA - Variáveis: PeerGroup_Refiners & Pipelines e R_ESG
 Estatística F: 22951.479356670523
 Valor p: 0.0

Anexo R- Modelo V: Variáveis mais significativas com base na significância individual

OLS Regression Results						
	coef	std err	t	P> t	[0.025	0.975]
Dep. Variable:	Sus_ESG		R-squared:	0.971		
Model:	OLS		Adj. R-squared:	0.971		
Method:	Least Squares		F-statistic:	1.053e+04		
Date:	Wed, 12 Jul 2023		Prob (F-statistic):	0.00		
Time:	18:31:49		Log-Likelihood:	-3374.2		
No. Observations:	1579		AIC:	6760.		
Df Residuals:	1573		BIC:	6793.		
Df Model:	5					
Covariance Type:	nonrobust					
const	-13.2539	0.406	-32.656	0.000	-14.050	-12.458
Sus_E	0.4695	0.003	141.661	0.000	0.463	0.476
Sus_S	0.4606	0.004	102.798	0.000	0.452	0.469
Sus_G	0.2686	0.004	64.877	0.000	0.261	0.277
Region_Europe	-0.2130	0.119	-1.793	0.073	-0.446	0.020
PeerGroup_Oil & Gas Producers	-0.1262	0.322	-0.392	0.695	-0.758	0.505
Omnibus:	3643.791		Durbin-Watson:	1.981		
Prob(Omnibus):	0.000		Jarque-Bera (JB):	19121980.376		
Skew:	-21.568		Prob(JB):	0.00		
Kurtosis:	540.386		Cond. No.	1.03e+03		

OLS Regression Results

Dep. Variable:	R_ESG	R-squared:	0.311
Model:	OLS	Adj. R-squared:	0.308
Method:	Least Squares	F-statistic:	141.7
Date:	Wed, 12 Jul 2023	Prob (F-statistic):	2.89e-124
Time:	18:28:24	Log-Likelihood:	-6467.8
No. Observations:	1579	AIC:	1.295e+04
Df Residuals:	1573	BIC:	1.298e+04
Df Model:	5		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
const	26.5790	1.728	15.384	0.000	23.190	29.968
R_E	0.1549	0.031	4.918	0.000	0.093	0.217
R_S	0.1484	0.035	4.246	0.000	0.080	0.217
R_G	0.2648	0.023	11.634	0.000	0.220	0.309
Region_Europe	4.1943	1.037	4.044	0.000	2.160	6.228
Region_North America	0.3144	0.904	0.348	0.728	-1.459	2.088

Omnibus:	141.537	Durbin-Watson:	1.821
Prob(Omnibus):	0.000	Jarque-Bera (JB):	207.619
Skew:	-0.687	Prob(JB):	8.24e-46
Kurtosis:	4.127	Cond. No.	586.

Anexo S- Multicolinearidade

Variável	VIF
Sus_E	0,993434
Sus_S	1,406597
Sus_G	1,420130
DivYield	1,055026
BookValue	1,069256
MarkCap	1,532729
Region_Australia	1,426924
Region_Europe	2,464483
Region_North America	2,689297
PeerGroup_Chemicals	1,143312
PeerGroup_Oil & Gas Producers	1,209907
PeerGroup_Pharmaceutical	1,153939

Fonte: Autor