



Facultad de Ciencias Económicas y Empresariales  
ICADE

**EVALUACIÓN Y ANÁLISIS DEL  
DESEMPEÑO MEDIOAMBIENTAL EN  
EMPRESAS ALTAMENTE  
CONTAMINANTES Y SU IMPACTO EN  
EL RENDIMIENTO FINANCIERO**

Autor: Ana Bueno Rodríguez  
Director: Dra. Elisa Aracil

MADRID | Junio 2024

## **RESUMEN**

Este trabajo de investigación tiene como objetivo analizar el impacto de las prácticas de gestión medioambiental en el desempeño financiero de empresas europeas pertenecientes a sectores con emisiones difíciles de reducir, tales como el sector del acero, el aluminio, el cemento y la industria química. Este estudio proporciona una perspectiva diferente al creciente interés académico y empresarial en este tema, enfocándose únicamente en las industrias altamente contaminantes. Se ha utilizado un panel de datos de 1.512 observaciones de 14 empresas europeas cotizadas, entre los años 2019 y 2022. Se han desarrollado modelos de regresión lineal para evaluar la relación entre las prácticas medioambientales y el rendimiento financiero, evaluado a través del ROA. Los resultados indican que la reducción de gases de efecto invernadero, una gestión eficiente del consumo eléctrico y una buena puntuación en gobernanza se correlacionan positivamente con el ROA. Este estudio proporciona una guía sobre cuáles son las variables medioambientales que tienen un mayor impacto en el desempeño financiero en las industrias de altas emisiones, y se destaca la necesidad de inversiones significativas en tecnologías renovables y la importancia de la financiación de transición.

***Palabras clave:** sectores difíciles de descarbonizar, gestión medioambiental, desempeño financiero, ROA, ESG, empresas europeas, emisiones, consumo energético, rentabilidad, aluminio, acero, química, cemento, sostenibilidad, financiación de transición.*

## **ABSTRACT**

This research aims to analyze the impact of environmental management practices on the financial performance of European companies in sectors with hard-to-reduce emissions, such as the steel, aluminum, cement, and chemical industries. This study provides a different perspective on the growing academic and business interest in this topic by focusing solely on highly polluting industries. A panel data set of 1,512 observations from 14 publicly traded European companies between the years 2019 and 2022 was used. Linear regression models were developed to evaluate the relationship between environmental practices and financial performance, assessed through ROA. The results indicate that the reduction of greenhouse gases, efficient management of electricity consumption, and good governance scores are positively correlated with ROA. This study provides a guide on which environmental variables have the greatest impact on financial performance in high-emission industries, highlighting the need for significant investments in renewable technologies and the importance of transition financing.

***Keywords:** hard-to-abate sectors, environmental management, financial performance, ROA, ESG, emissions, energy consumption, profitability, aluminum, steel, chemical, cement, sustainability, transition financing.*

## **TABLA DE CONTENIDOS:**

<b>RESUMEN.....</b>	<b>2</b>
<b>ABSTRACT.....</b>	<b>3</b>
<b>1. INTRODUCCIÓN .....</b>	<b>5</b>
<b>1.1. Objetivo.....</b>	<b>5</b>
<b>1.2. Pregunta de investigación.....</b>	<b>5</b>
<b>1.3. Justificación selección del tema .....</b>	<b>5</b>
<b>1.4. Respuestas y desafíos globales y europeos ante el cambio climático.....</b>	<b>6</b>
<b>1.5. Industria de los sectores con emisiones difíciles de reducir .....</b>	<b>8</b>
<b>1.6. Metodología propuesta y estructura .....</b>	<b>10</b>
<b>2. MARCO TEÓRICO: Relación entre las prácticas medioambientales y el desempeño financiero .....</b>	<b>12</b>
<b>3. MÉTODOS .....</b>	<b>16</b>
<b>3.1. Metodología .....</b>	<b>16</b>
<b>3.2. Muestra .....</b>	<b>16</b>
<b>3.3. Medidas.....</b>	<b>19</b>
<b>3.4. Modelo.....</b>	<b>24</b>
<b>4. RESULTADOS .....</b>	<b>27</b>
<b>4.1. ANÁLISIS DESCRIPTIVO .....</b>	<b>27</b>
<b>4.1.1. Distribución de las variables numéricas .....</b>	<b>28</b>
<b>4.1.2. Valores atípicos .....</b>	<b>33</b>
<b>4.1.3. Normalización de las variables .....</b>	<b>38</b>
<b>4.1.4. Análisis de correlación.....</b>	<b>44</b>
<b>4.2. MODELO REGRESIÓN LINEAL .....</b>	<b>46</b>
<b>4.2.1. Análisis de la multicolinealidad .....</b>	<b>46</b>
<b>4.2.2. Resultados obtenidos en cada modelo de regresión lineal.....</b>	<b>48</b>
<b>5. DISCUSIÓN, CONCLUSIONES Y LIMITACIONES.....</b>	<b>55</b>
<b>6. ANEXO .....</b>	<b>59</b>
<b>7. DECLARACIÓN DE USO DE HERRAMIENTAS DE INTELIGENCIA ARTIFICIAL GENERATIVA EN TRABAJOS DE FIN DE GRADO.....</b>	<b>66</b>
<b>8. BIBLIOGRAFÍA.....</b>	<b>68</b>

# 1. INTRODUCCIÓN

En esta primera sección del trabajo se define el objetivo general del trabajo y la pregunta de investigación que se va a resolver. Asimismo, se explica la importancia del objeto de estudio. Finalmente, se incluye la estructura seguida.

## 1.1. Objetivo

El objetivo de este trabajo de investigación se centra en analizar el impacto que tendrían diversas prácticas de gestión medioambiental en el desempeño financiero de empresas europeas que operan en sectores con emisiones difíciles de reducir.

## 1.2. Pregunta de investigación

La pregunta de investigación que se va a responder con este trabajo es: ¿Cómo impactan las diferentes actividades y prácticas de gestión medioambiental y las puntuaciones ESG en el rendimiento financiero de las empresas europeas que operan en sectores con emisiones difíciles de reducir?

## 1.3. Justificación selección del tema

Cada vez es mayor el interés tanto académico como empresarial que existe en comprender cómo las prácticas de gestión medioambiental impactan en el rendimiento financiero de las empresas. Numerosos estudios sugieren que las empresas con mejor prácticas ESG, pueden obtener unos resultados financiero más rentables. Aproximadamente el 90% de los estudios encuentran una relación no negativa entre la gestión sostenible y el rendimiento financiero, y la mayoría reporta resultados positivos (Bush, Bassen & Friede, 2015).

Es cierto que numerosos académicos como Freeman (1984), Hart (1995), Elkington (1997), Porter & Kramer (2011) y Busch & Hoffmann (2011), han estudiado la relación entre la sostenibilidad y el rendimiento financiero. Sin embargo, existen limitados estudios que se centren únicamente en los sectores con emisiones difíciles de reducir. Estudios como el de Porter (1997), Michael et al. (2007), Clark (2015), Freiner (2015), y Viehs (2015) o el de Alshehhi, Khare y Nobanee (2018) analizan cómo las practicas sostenibles impactan en el rendimiento financiero de diferentes industrias, pero no se centran exclusivamente en los sectores con emisiones difíciles de reducir.

Por tanto, este trabajo de investigación beneficiará a la literatura existente, ofreciendo nuevos conocimientos y análisis. El estudio tiene como objetivo proporcionar datos empíricos y evidencias adicionales y actuales, únicamente centralizado en la industria de los sectores de emisiones difíciles de reducir, concretamente el sector del acero, el aluminio, la química y el cemento.

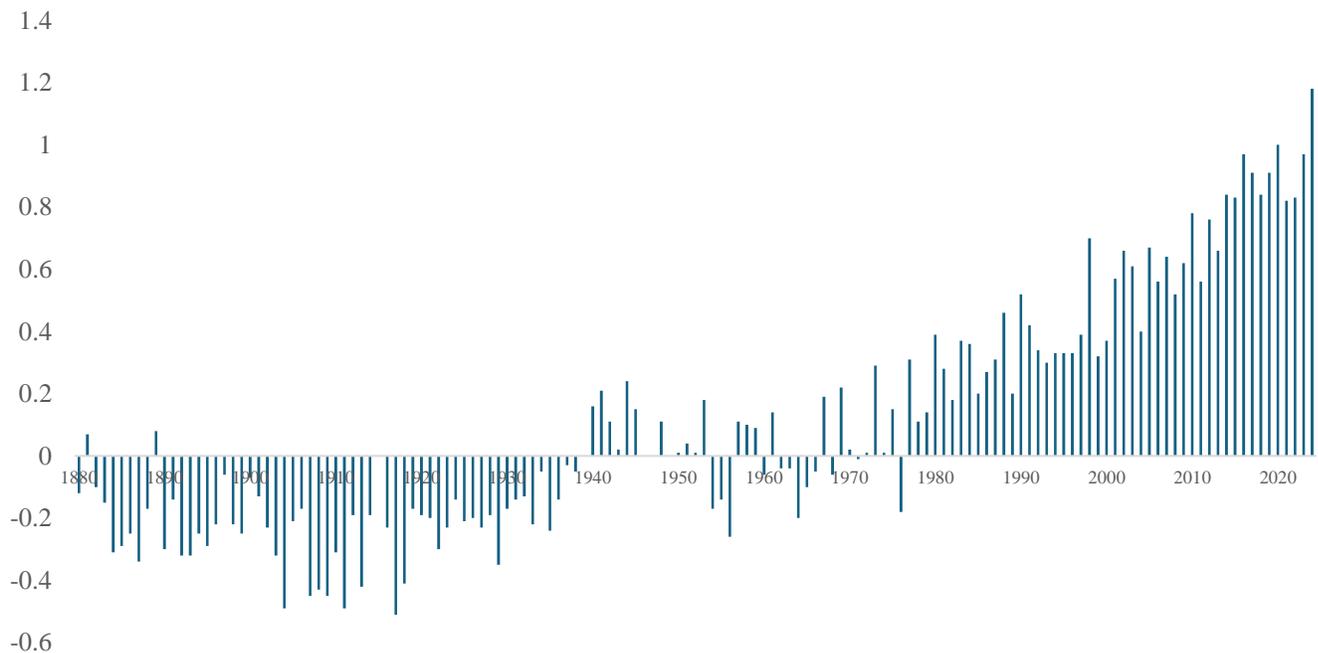
Proporcionar estudios actuales es crucial dentro de la naturaleza dinámica de la sostenibilidad y la regulación sostenible para poder responder de forma eficaz a los desafíos y oportunidades que surgen en este campo. Asimismo, este estudio va a ofrecer una perspectiva cuantitativa amplia. Se va a proveer un análisis detallado y multifacético del impacto de diferentes variables cuantitativas como son las emisiones de dióxido de carbono (CO<sub>2</sub>), emisiones de gases de efecto invernadero (GEI), consumo de agua, cantidad de recursos, uso de la electricidad y empleo de energía renovable, así como distintas puntuaciones ESG que abarquen diferentes aspectos de la sostenibilidad en el rendimiento financiero de estas industrias.

#### **1.4. Respuestas y desafíos globales y europeos ante el cambio climático**

El cambio climático se considera uno de los mayores riesgos ambientales para la sociedad. La Convención Marco de Naciones Unidas sobre el Cambio Climático, a través del Acuerdo de París firmado en 2015, busca limitar el aumento de la temperatura global a menos de 2 grados Celsius (Herzog, Kheshgi, Morri & Paltsev, 2021).

En el gráfico 1 se presenta la evolución de la temperatura media global (°C) desde 1880 hasta el 2024. Se puede observar cómo ha ido incrementando progresivamente desde el año 1980.

*Gráfico 1. Evolución de la temperatura media global (°C) 1880-2024*



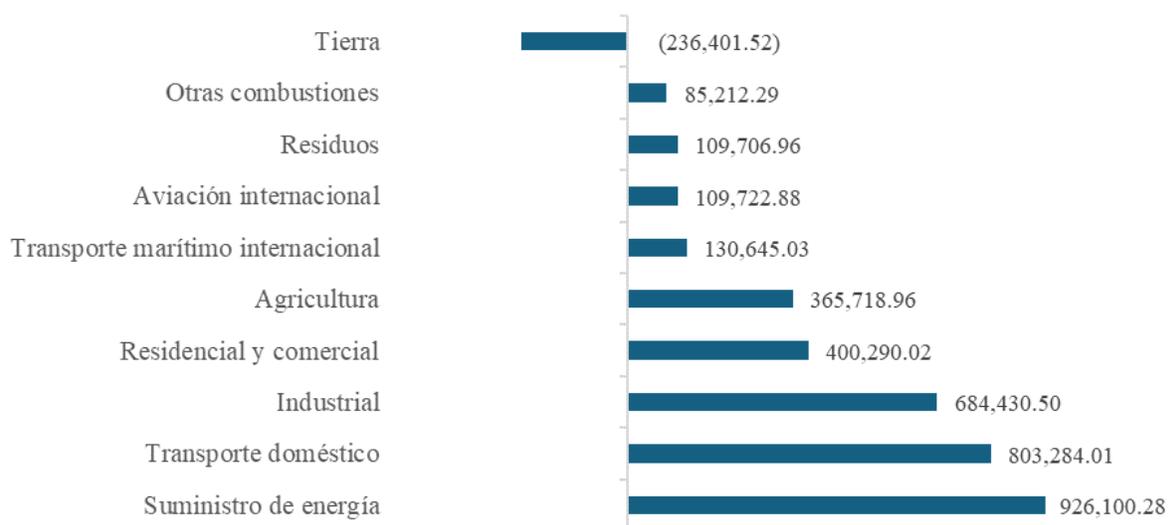
*Fuente: NASA's Goddard Institute for Space Studies (GISS).*

La Unión Europea (UE) ha desempeñado un papel destacado como potencia líder en la transformación hacia una economía más sostenible desde el punto de vista energético y ecológico. El Pacto Verde Europeo representa el hito más reciente en este proceso, integrando una serie de iniciativas y marcos legales con el objetivo de abordar los desafíos climáticos y ambientales. La visión del Pacto Verde se centra en lograr una economía eficiente sin emisiones netas de Gases de Efecto Invernadero (GEI) para el año 2050, con una propuesta de reducción entre el 50% y 55% para el año 2030, en comparación con las emisiones de 1990 (Comisión Europea, 2019).

Las emisiones antropogénicas de gases de efecto invernadero (GEI) deben reducirse en todos los sectores. En 2022, aproximadamente el 20% de las emisiones de CO<sub>2</sub> del sistema energético global se originaron en el sector industrial. Para limitar el aumento de la temperatura global a 2° C o menos, todas las industrias intensivas en emisiones, incluyendo la industria del acero, la industria del cemento y la industria química, deben lograr emisiones netas cero hacia el 2080 (Ahman et al., 2017).

En el gráfico 2 se observa la cantidad de gases de efecto invernadero medida en kt CO eq generada por cada uno de los sectores en la Unión Europea en el año 2022. Se puede verificar que aproximadamente un 20% corresponde al sector industrial.

*Gráfico 2. Participación sectorial en la Unión Europea de GEI 2022 (kt CO eq)*



*Fuente: Agencia Europea de Medio Ambiente*

### **1.5. Industria de los sectores con emisiones difíciles de reducir**

Los sectores de industrias con emisiones difíciles de reducir engloban a todos aquellos en los que la reducción de emisiones de gases de efecto invernadero es particularmente desafiante debido a varias razones técnicas, económicas y estructurales, como son la industria del aluminio, el acero, la química o el cemento. Estas industrias generalmente tienen procesos que dependen en gran medida de la energía intensiva y, a menudo, requieren temperaturas extremadamente altas o procesos químicos específicos que actualmente son difíciles de descarbonizar. De ahí, que la reducción de emisiones en estos sectores sea uno de los principales desafíos.

Distintas fuentes de datos ofrecen estimaciones variadas sobre las emisiones globales de gases de efecto invernadero (GEI) causadas por estas industrias. La industria de los sectores de emisiones difíciles de reducir contribuyó al 33% de las emisiones mundiales del GEI (Risman et al., 2014). Las industrias pesadas como cemento, hierro, acero y productos químicos son industrias que

generan de alrededor del 22% de las emisiones mundiales de CO<sub>2</sub> (Bataille, 2019). Actualmente, estas emisiones continúan aumentando debido a la creciente demanda de materiales básicos.

El Informe World Energy Outlook 2022 de la Agencia Internacional de la Energía, reporta que si no se toman medidas en dichas industrias las temperaturas aumentarían una media de 2,5°C, lo que implicaría que el calentamiento global no se podrá contener en un 1,5°C. Por eso, la única forma de mantener el calentamiento en unos niveles manejables es si se logra conseguir unas emisiones netas cero para el 2050. Sin embargo, tanto el Panel Intergubernamental de Expertos sobre cambio Climático (IPCC) como la Comisión Europea reconocen que sin la captura y la descarbonización de dichas industrias no es posible alcanzar los objetivos de reducción para 2030 y especialmente para 2050.

Alinear la trayectoria de la reducción de las emisiones con los objetivos del Acuerdo de París es una tarea urgente y difícil. Estos sectores son una parte fundamental de la sociedad industrializada y han evolucionado a lo largo de varias décadas con las infraestructuras, las instituciones sociales, el conocimiento y la tecnología. Esto ha conducido a un sistema de producción eficiente, pero dependiente de los combustibles fósiles. El uso de energía y materia prima fósil está profundamente incrustado en el proceso de producción. Ejemplos incluyen el uso de coque para reducir el mineral de hierro en altos hornos, crudo de petróleo como materia principal para los petroquímicos y gas natural como materia prima para el amoníaco. La dependencia puede desbloquearse, pero para ello es necesario un esfuerzo concertado y una visión a largo plazo. Se necesitan desarrollar y cambiar a nuevas tecnologías de proceso innovadoras que utilicen energía y materias primas 100% renovables.

Sin embargo, hoy en día, en todos estos sectores la transición a prácticas más sostenibles es compleja debido a barreras tecnológicas y costes prohibitivos. La descarbonización de sectores como en el de la producción de acero y cemento requieren inversiones significativas en innovaciones tecnológicas, como la implementación de sistemas de captura, utilización y almacenamiento de carbono (CCUS) o la transición a procesos y combustibles de bajas emisiones, como el hidrógeno. La industria química enfrenta costes iniciales y contiene una diversidad de productos que dificultan la transición a tecnologías limpias. En el transporte pesado, la

dependencia de combustibles fósiles y la complejidad logística obstaculizan la descarbonización. La producción de cemento, altamente dependiente de combustibles fósiles, enfrenta limitaciones en la búsqueda de alternativas sostenibles debido a la demanda global constante y la complejidad de los procesos químicos.

La introducción de las distintas innovaciones medioambientales y cambios de vías de producción sin emisiones tienen un coste: Se espera que el acero sea entre un 20% y un 40% más caro si se produce a partir de hidrógeno renovable (Vogl et al. 2018), el cemento podría costar entre un 70% y un 100% más con el concreto autocompactante (CAC) (Johnsson & Rootzén, 2016) y los plásticos podrían ser un 200% más caros si se produjeran con electricidad renovable e hidrógeno en lugar de materias primas fósiles de bajo coste (Palm et al., 2016). No obstante, este incremento de costes de producción representa un aumento de al menos entre el 1% y el 2% del precio de venta final de por ejemplo un coche o un conjunto de apartamentos (Rootzén, 2015).

Por todo ello, el propósito de este trabajo se centra en analizar el impacto financiero que tendrían la implementación de distintas prácticas de gestión medioambiental en empresas que operan en sectores con emisiones difíciles de reducir.

### **1.6. Metodología propuesta y estructura**

El procedimiento que será llevado a cabo para la realización de este trabajo es el siguiente: En primer lugar, se comenzará con un estudio de investigación descriptiva del marco teórico, realizando un análisis de la literatura. Para ello, se obtendrá la información de artículos relevantes encontrados en Google Scholar. En esta primera etapa de investigación, se abarcará el impacto de la actividad medioambiental en el desempeño financiero de las empresas y se formularán las hipótesis que serán objeto de estudio. Tras ello, se continuará con el análisis enfocándose en los métodos de investigación empleados. Se examinará tanto la metodología, como la base de datos, variables y modelos utilizados. Posteriormente, se presentarán los resultados de los modelos con un previo análisis descriptivo. Finalmente, se proporcionarán unas conclusiones y limitaciones del trabajo.

El trabajo se estructura en cuatro partes. El primer bloque presenta el marco teórico. El segundo bloque recoge todo el análisis del conjunto de datos y la metodología empleada. En un tercer

apartado, se presentan los distintos resultados. Por último, se facilitan unas conclusiones y limitaciones del trabajo.

## **2. MARCO TEÓRICO: Relación entre las prácticas medioambientales y el desempeño financiero**

Numerosos autores han explorado cómo el desempeño económico de las empresas se ve afectado por la implementación de cambios de la actividad medioambiental, constituyendo así un tema ampliamente estudiado en la literatura académica. Sin embargo, este trabajo de investigación contribuirá a la literatura existente aportando nuevos conocimientos, sometiendo a prueba y contrastando diversas hipótesis que están centradas exclusivamente en la industria de los sectores difíciles de descarbonizar.

Entre los siglos XVIII y XX, gracias al desarrollo industrial se experimentó un gran crecimiento económico (Garcés-Ayerbe, Murillo-Luna & Rivera-Torres, 2007; Murillo, 2005). En consecuencia, del desarrollo industrial, el planeta ha sufrido una degradación medioambiental incontrolable (Bahillo, Grábalos & Lafuente, 2007; Castillo-González 2018; Shrivastava, 1995), dando lugar a lo conocido como cambio climático. Gran parte de este deterioro ha sido causado por la actividad empresarial que ha elevado las emisiones de gases de efecto invernadero y el uso excesivo e incontrolable de recursos (Moors, Mulder & Vergragt, 2005). Existen numerosos datos proporcionados por la ONG de Manos Unidas que sustentan estas afirmaciones y evidencian la existencia de ese daño medioambiental. Desde el siglo XX, el mundo industrializado ha aumentado en un 30% la concentración de gases de efecto invernadero. La temperatura media global entre 1880 y 2012 ha incrementado 0.85 grados. Asimismo, desde los años 70 se han observado muchas regiones como en Sahel, el mediterráneo, África meridional y partes de Asia meridional con sequías intensas y prolongadas. De ahí, que en siglo XXI se haya incrementado la presión sobre las empresas para que adopten medidas con el objetivo de minimizar su huella medioambiental.

Desde la adopción de los Objetivos de Desarrollo Sostenible en 2015, ha habido un interés creciente entre el impacto de la aplicación de estas prácticas medioambientales y el desempeño financiero de las empresas (Vorontsova et al., 2022). En los últimos años, se ha intensificado el estudio de la relación entre las prácticas sostenibles y el desempeño financiero (Betti et al., 2018). El propósito de las empresas se orienta en alcanzar un equilibrio entre los factores sociales y medioambientales y su desempeño financiero.

Autores como Porter (1991), Sells & Jankousky (1994), Hart (1995), Porter & Van der Linde (1995) o Xepapadaeas & Zeeuw (1999), han justificado unos resultados económicos más consistentes y unas ventajas competitivas para aquellas empresas que invierten en tecnologías y en prácticas que reducen sus emisiones de efecto invernadero. Es decir, sugieren que es posible alcanzar los objetivos de calidad medioambiental a la vez que se reducen los costes de producción o se incrementa la productividad. Esta justificación planteada por estos autores queda denominada en lo conocido como “Hipótesis Porter”. La tesis de Porter (1991) afirmaba que las empresas que reducen sus emisiones en la atmósfera son las que mayor eficiencia productiva demuestran fue confirmada por Burnelt & Hansen (2008), quienes investigaron la relación entre las emisiones CO2 y la eficiencia productiva de las empresas del sector eléctrico de Estados Unidos.

- **Hipótesis 1:** una reducción significativa de las emisiones de gases de efecto invernadero lleva a una mejora en el desempeño financiero.

Distintos autores como Penrose (1959) y Wernerfelt (1984) a través de “La Teoría de Recursos y Capacidades” han examinado la forma en la que las inversiones medioambientales proactivas mejoran la competitividad empresarial. Esta teoría propone que la ventaja competitiva únicamente se puede sostener mediante recursos y capacidades que sean únicos y socialmente complejos, lo que impide su fácil replicación por parte de los competidores (Barney & Grant, 1991; Rumelt, 1984). Dentro de este campo, uno de los trabajos más relevantes es el de “Teoría de recursos naturales de la empresa” realizado por Hart (1995). En él, se argumenta que en un contexto donde las prácticas económicas actuales no son ambientalmente sostenibles, la ventaja competitiva se derivará cada vez más de habilidades específicas como la reducción de residuos o el diseño de productos ecológicos.

- **Hipótesis 2:** la implementación de sistemas de gestión de recursos naturales, como la optimización del consumo energético, se traduce en un mayor desempeño financiero.
- **Hipótesis 3:** un aumento en el uso de energía renovables lleva a un mayor desempeño financiero.

Con el tiempo, este estudio ha sido profundizado por otros autores, y se ha evidenciado que el desempeño financiero es mayor con la implementación de recursos voluntarios como la cultura

organizativa, la coordinación entre las diferentes funciones de la empresa y la influencia política (Fouts & Russo, 1997), así como los activos tecnológicos (Dowell, Hart & Yeung, 2000; Klassen & McLaughlin, 1996; Klassen & Whybark, 1999; Rugman & Verbeke, 1998). También se ha enfatizado la importancia de la actitud de los directivos (Sharma, 2000), el aprendizaje organizacional (Sharma & Vredenburg, 1998), la capacidad de innovar de manera continua y de integrar a los *stakeholders* en los procesos empresariales (Fouts & Russo, 1997; Sharma & Vredenburg, 1998), la incorporación de consideraciones ambientales en la planificación estratégica (Douglas & Judge, 1998) y el aprovechamiento de las sinergias o complementariedades (Christmann, 2000; Rugman & Verbeke, 1998).

Por todo ello, la literatura sugiere que una buena gestión de los aspectos ESG (Ambientales, Sociales y de Gobernanza) pueden impactar positivamente en el rendimiento financiero (Clark, Feiner, & Viehs, 2015; Eccles, Ioannou, & Serafeim, 2014). Altas puntuaciones ESG reflejan prácticas sostenibles efectivas y una mejor gestión de riesgos, lo cual se traduce en mejores resultados financieros. Gompers, Ishii & Metrick (2003) destacan que una buena gobernanza corporativa puede mejorar la eficiencia y reducir los riesgos, lo que se traduce en un mejor rendimiento financiero. Su estudio “*Corporate Governance and Equity Prices*” afirma que las empresas con mayores puntuaciones en sostenibilidad y gobernanza tienden a lograr una mayor valorización de mercado y rentabilidad. Asimismo, esto está sustentado también por un informe de McKinsey & Company (2020) que destaca que las empresas líderes en prácticas ESG tienden a lograr un coste de capital más bajo y una mayor valoración en el mercado.

- **Hipótesis 4:** una mayor puntuación en los distintos componentes de ESG tienen un mayor desempeño financiero.

Sin embargo, para contribuir a la literatura de manera significativa y aportar una perspectiva nueva, es importante resaltar que las hipótesis definidas en esta investigación se han examinado específicamente en empresas europeas pertenecientes a industrias con niveles de emisiones altamente contaminantes y tratan de completar la literatura existente en este tema a través de cuatro modelos de regresión lineal y la combinación de diferentes variables financieras como son el ROA o el EBIT ratio e indicadores medioambientales como la cantidad de emisiones de gases efecto

invernaderos y dióxido de carbono de alcances 1 y 2, uso de la electricidad, cantidad de agua consumida, número de derrames y de multas medioambientales y distintas puntuaciones ESG, entre otros.

### **3. MÉTODOS**

#### **3.1. Metodología**

Para llevar a cabo una investigación exhaustiva de las hipótesis previamente planteadas y determinar y evaluar su validez en el conjunto de datos estudiados, se realizará un modelo predictivo. El modelo predictivo elegido será complementado con un previo análisis tipo descriptivo que permitirá observar las características fundamentales y detectar los patrones existentes dentro del conjunto de datos. Para ello, se aplicarán cuatro enfoques de estudio distintos, que incluyen la exploración de la distribución de las variables numéricas, la identificación de valores atípicos, la normalización de variables y un análisis de correlación entre las variables.

Una vez se hayan trabajado las variables, se elaborará un modelo predictivo con el fin de estimar el desempeño financiero de las empresas basándose en diferentes variables financieras y medioambientales. Existen diversos modelos predictivos, para el estudio de este trabajo se va a trabajar con un modelo de regresión lineal. Esta sección incluirá una explicación teórica sobre la importancia del análisis predictivo y la implementación del modelo de regresión lineal.

Ambos análisis se realizarán utilizando en el entorno de software de R Studio. Al final del documento se muestra el código que ha sido utilizado.

#### **3.2. Muestra**

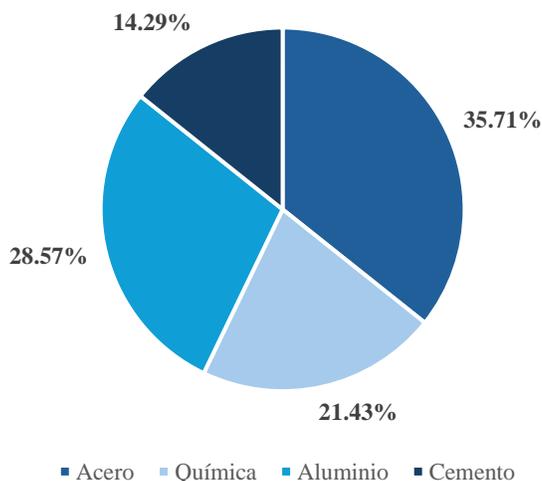
Se va a emplear un panel que cuenta con 1.512 observaciones correspondientes a 14 empresas europeas que han cotizado en bolsa en el periodo 2019-2022. Todas ellas pertenecen a las cuatro industrias de los sectores altamente contaminantes: acero (35.71%), aluminio (28.57%), química (21.42%) y cemento (14.29%) de distintos países europeos.

El criterio que respalda la selección de esta muestra es que todas ellas son empresas representativas del sector así que proporcionan una visión completa de los desafíos a los que se enfrentan cada una de las industrias con emisiones difíciles de reducir. Asimismo, son compañías con un gran tamaño y alcance, por tanto, sus acciones tienen un impacto significativo en la reducción de emisiones a nivel global. Por otro lado, varias de las empresas seleccionadas como es el caso de SSAB con la tecnología Hybrit o BASF con los proyectos de hidrógeno verde, son pioneras en la

adopción de tecnologías avanzadas para la reducción de emisiones. Así que observar el impacto económico de las prácticas sostenibles de estas empresas permite evaluar cómo la sostenibilidad puede ser compatible con la rentabilidad y la competitividad, sirviendo de modelo para empresas similares en este sector. Finalmente, son empresas que cotizan en bolsa y publican informes de sostenibilidad y medioambientales. Toda esta disponibilidad y acceso a datos relevantes, proporciona una sólida base de información para el trabajo de investigación. Por todo ello, la elección de estas empresas permite realizar un análisis detallado de cómo las prácticas medioambientales afectan al rendimiento financiero de empresas en sectores con emisiones difíciles de reducir.

En la figura 1 se muestra la proporción de cada uno de los sectores que hay en la muestra. La industria del acero es la más representada (35.71%), mientras que hay una menor cantidad de empresas dentro del sector del cemento (14.29%).

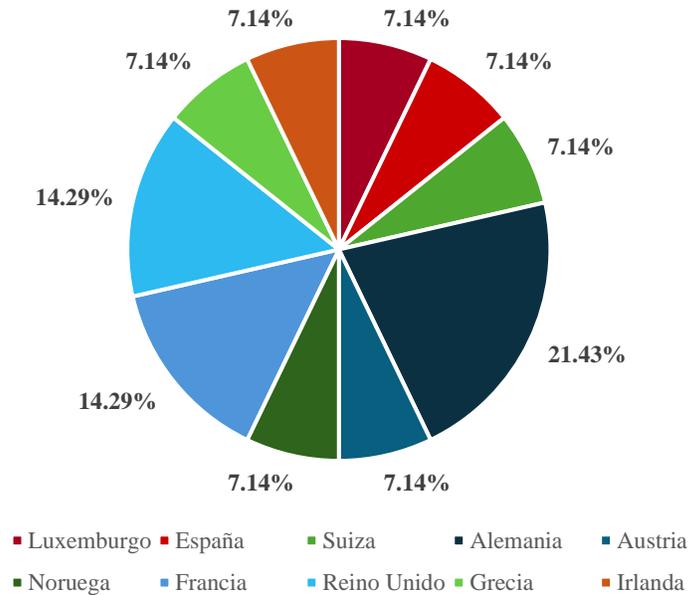
*Figura 1. Representación de la muestra por sectores*



*Fuente: Elaboración propia*

Por su parte, en la figura 2 se representan el total de empresas de la muestra divididas por geografía. Todas ellas son europeas, sin embargo, se puede ver que no todos los países contribuyen en la misma manera. Alemania es el país al que pertenecen una mayor cantidad de empresas (21.43%).

Figura 2. Representación de la muestra por geografía



*Fuente: Elaboración propia*

La utilización de la metodología de datos de panel se justifica por la muestra, la cual combina series de tiempo por 4 años (2019-2022) y datos de sección cruzada para 14 empresas, lo que nos permite hacer un uso óptimo de dicha metodología (Arellano, 2003). Esta metodología de datos panel ofrece beneficios al permitir manejar ciertas dificultades asociadas con la heterogeneidad constante e inobservable (Arellano & Bover, 1990). Esto se refiere a atributos particulares de cada empresa que se mantienen constante en el tiempo y que serán representadas como un término de efectos fijos en el análisis.

En la tabla 1 se recoge información relevante de las empresas, que han sido escogidas como objeto de estudio:

Tabla 1. Características de la muestra

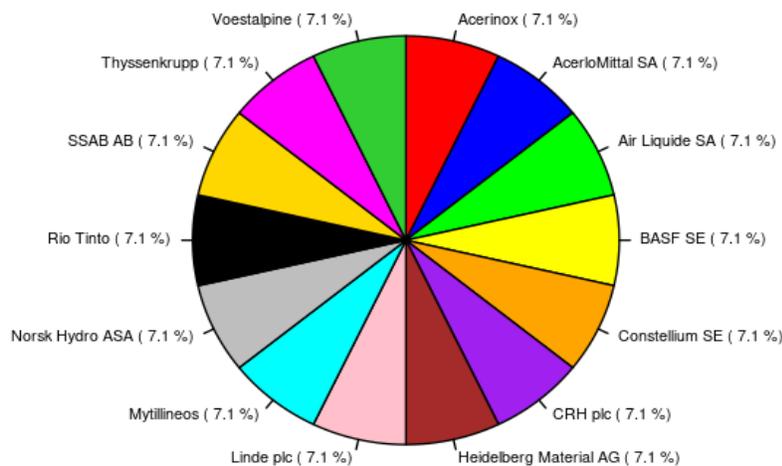
	AcerloMittal SA	Acerinox	SSAB AB	Thyssenkrupp	Voestalpine	Norsk Hydro ASA	Constellium SE
<b>Acrónimo</b>	MTS.MC	ANIOY	SSAAY	TKA.DE	VOE.VI	NHYKF	CSTM
<b>Sector</b>	Acero	Acero	Acero	Acero	Acero	Aluminio	Aluminio
<b>Año de fundación</b>	2006	1970	1978	1999	1938	1905	2011
<b>País de la Sede</b>	Luxemburgo	España	Suiza	Alemania	Austria	Noruega	Francia
<b>Número de empleados</b>	154,352.00	8,298.00	14,632.00	99,981.00	51,212.00	32,014.00	12,500.00
<b>Market Cap (Junio 20)</b>	18.473bn	2.703bn	5.892bn	2.619bn	4.492bn	13.907bn	2.998bn
<b>Volumen</b>	163,189	8,200	1,220	6,550	150	5,801	740,629
<b>Precio Acción</b>	25.3	5.41	3.355	5.27	27.69	6.19	22.43
<b>Beta(5 años al mes)</b>	1.88	1.25	1.17	1.97	1.40	1.34	1.99

	Rio Tinto	Mytilineos	BASF SE	Air Liquide SA	Linde plc	Heidelberg Material AG	CRH plc
<b>Acrónimo</b>	RIO	MYTHY	BAS	AIQUY	LIN	HDELY	CRH
<b>Sector</b>	Aluminio	Aluminio	Química	Química	Química	Cemento	Cemento
<b>Año de fundación</b>	1995	1990	1990	1902	2018	1873	1970
<b>País de la Sede</b>	Reino Unido	Grecia	Alemania	Francia	Reino Unido	Alemania	Irlanda
<b>Número de empleados</b>	54,000.00	3,444.00	111,315.00	67,100.00	60,422.00	75,800.00	78,500.00
<b>Market Cap (Junio 20)</b>	109.259bn	5.566bn	41.547bn	116.335bn	208.712bn	19.571bn	55.433bn
<b>Volumen</b>	3,887,030	212,361	226,640	98,442	1,546,331	32,755	2,769,255
<b>Precio Acción</b>	63	36.5	52.56	40.58	453	20.67	83.46
<b>Beta(5 años al mes)</b>	0.65	0.91	1.26	0.66	0.94	1.33	1.35

Fuente: Yahoo Finance

Asimismo, en la figura 3, se observa que la distribución de las empresas presentes en el conjunto de datos es equitativa y homogénea.

Figura 3. Representación de la muestra



Fuente: Elaboración propia (Rstudio)

### 3.3. Medidas

Las fuentes de datos de estas variables han sido muy variadas. Todas ellas son empresas que cotizan en bolsa, por tanto, los indicadores financieros se han podido obtener de Capital IQ. Por su parte, los indicadores medioambientales se han sacado tanto de los informes de sostenibilidad

de cada una de las empresas en los años correspondientes como de Bloomberg que es una plataforma muy relevante en el ámbito financiero.

A continuación, se muestra una definición de cada una de las variables que constituyen la base de datos, junto con su acrónimo, en los casos que aplica. Además, en las variables medioambientales su fuente correspondiente:

### ***Variable dependiente***

- ROA (Return on Assets): porcentaje que mide la eficiencia de la empresa en la utilización de sus activos para generar beneficios.

En este estudio para poder evaluar el desempeño financiero, se ha escogido el ROA como la variable dependiente. El ROA permite a las empresas evaluar su rentabilidad y compararla con la competencia. Asimismo, les brinda la oportunidad de poder identificar sus áreas de mejora, optimizando su desempeño financiero y aumentando la competitividad en el mercado (Cox et al., 2014). Además, estos modelos sirven como una herramienta a largo plazo de planificación financiera. Les proporcionan estimaciones del ROA que permite a las compañías establecer metas realistas y proyectar ganancias (Demyantk et al., 2010). Finalmente, pueden ayudar a identificar y mitigar los posibles riesgos operativos y financieros, salvaguardando los activos del banco (Petropoulos et al., 2017).

### ***Variables independientes incluidas en los modelos***

- EBIT: ganancias de la empresa antes de deducir los intereses e impuestos, reflejando únicamente su rendimiento operativo. Medido en millones de dólares (USD).
- Ratio EBIT (EBIT ratio): medida que sirve para comparar el EBIT con los ingresos, reflejando que porcentaje de los ingresos se convierte en ganancias antes de intereses e impuestos.

- % Alcances 1 y 2 (Objetivo1y2): porcentaje que representa la cantidad de emisiones directas e indirectas de gases de efecto invernadero que cada una de las empresas se ha fijado como objetivo para reducir en un número de años. *Fuente: Informes de sostenibilidad de cada una de las empresas (2019-2022).*
- Año Inicial Alcances 1 y 2 (AnoInicio1y2): año en el que cada una de las empresas ha establecido el porcentaje objetivo a reducir de los alcances 1 y 2 de emisiones de gases de efecto invernadero. *Fuente: Informes de sostenibilidad de cada una de las empresas (2019-2022).*
- Año Final Alcances 1 y 2 (AnoFin1y2): año meta para el cumplimiento de reducción del porcentaje objetivo de los alcances 1 y 2 de emisiones de gases de efecto invernadero por cada una de las empresas. *Fuente: Informes de sostenibilidad de cada una de las empresas (2019-2022).*
- Residuos totales (Residuos): volumen total de desechos, tanto peligrosos como no peligrosos, generados por una empresa. Variable medida en miles de toneladas métricas. *Fuente: Bloomberg.*
- Uso de energía Renovable (UsoEnergiaRenovable): cantidad de energía consumida por una empresa que proviene de fuentes renovables. Esta variable incluye la compra de Certificados de Energía Renovable (REC) por parte de las empresas. Se mide en megavatios hora (MWh). *Fuente: Bloomberg.*
- Electricidad Utilizada (UsoElectricidad): cantidad total de electricidad consumida por una empresa en un periodo de tiempo. Se mide en miles de megavatios hora (MWh). *Fuente: Bloomberg.*
- Número de Derrames (Derrames): número total de derrames de materiales peligrosos generados la empresa en un determinado período. *Fuente: Bloomberg.*

- Uso total de agua (*UsoAgua*): cantidad total de agua utilizada por una compañía para sus procesos operativos, Esta medido en miles de metros cúbicos. *Fuente: Bloomberg.*
- Número de multas ambientales (*MultasAmbientales*): número de multas ambientales pagadas por la empresa cada año, que han sido impuestas por el incumplimiento de las leyes y regulaciones ambientales. *Fuente: Bloomberg.*
- Alcance de CO2 (dióxido de carbono) 1 y 2 basado en ubicación (*CO2 1y2*): emisiones totales de dióxido de carbono generadas por empresa, distinguiendo entre las emisiones directas (Alcance 1) y las emisiones indirectas (Alcance 2). Esta variable se mide en miles de toneladas métricas y ha sido calculada utilizando el método de contabilidad basado en la ubicación. *Fuente: Bloomberg.*
- Alcance de GEI (Gases de Efecto Invernadero) alcances 1 y 2 basados en ubicación (*GEI 1y2*): emisiones directas e indirectas de gases de efecto invernadero generadas por una empresa. Se definen como gases de efecto invernadero aquellos gases que contribuyen a la retención del calor en la atmósfera terrestre, incluyendo el dióxido de carbono (CO2), metano, óxido nitroso y otros. Esta variable se mide en miles de toneladas métricas y se ha utilizado el método de contabilidad basado en la ubicación. *Fuente: Bloomberg.*
- Puntuación de riesgo climático gubernamental (*CLIMATE SCORE*): proporciona una puntuación que mide el progreso del país o región del gobierno al que pertenece la empresa en cumplir con los objetivos establecidos en el Acuerdo de Paris. La puntuación varía de 0 a 10; siendo el 10 la mejor. *Fuente: Bloomberg.*
- Puntuación del cambio climático del Proyecto de Divulgación de carbono (*CDP CLIMATE CHANGE SCORE*): es la calificación que una empresa recibe en su informe de cambio climático por el proyecto de divulgación de carbono. Esta calificación mide la transparencia y la calidad de la información que las empresas proporcionan sobre sus impactos, riesgos y oportunidades relacionados con el cambio climático. La puntuación

oscila entre el 1 y el 10. Una puntuación cercana a 10 indica un buen desempeño en la gestión y divulgación relacionada con el cambio climático. *Fuente: Bloomberg.*

- Ranking ASG de la S&P (SP ESG RANK): es el ranking ASG (Ambiental, Social y Gubernamental) proporcionado por S&P Global, en el que se evalúa el desempeño de una empresa en términos de criterios ambientales, globales y de gobernanza. La calificación va del 1 al 100, donde un mayor más elevado indica un mejor desempeño en las prácticas ASG. *Fuente: Bloomberg.*
  
- Puntuación ASG (ESG SCORE): es la puntuación de Bloomberg que evalúa el desempeño agregado de Gobernanza, Social y Ambiental (ASG) de la empresa. La puntuación se basa en la visión de Bloomberg de la materialidad financiera ESG. Se calcula mediante una media generalizada ponderada de las puntuaciones de pilares, donde los pesos están determinados por el ranking de prioridad de los pilares. Los valores varían de 0 a 10, siendo el 10 el mejor. *Fuente: Bloomberg.*
  
- Puntuación del Pilar Ambiental: (ENVIRON SCORE): proporciona la puntuación de Bloomberg que evalúa el desempeño ambiental agregado de la empresa. Los valores oscilan entre 0 y 10; siendo 10 el mejor. *Fuente: Bloomberg.*
  
- Puntuación del Pilar de Gobernanza (GOV SCORE): proporciona la puntuación de Bloomberg que evalúa el desempeño de gobernanza agregado de la empresa. Los valores oscilan entre 0 y 10; siendo 10 el mejor. *Fuente: Bloomberg.*
  
- Puntuación del Pilar Social (SOCIAL SCORE): calificación de Bloomberg que evalúa las relaciones sociales que tiene una empresa con los empleados, proveedores, clientes y comunidades donde opera. Los valores oscilan entre 0 y 10; siendo 10 el mejor. *Fuente: Bloomberg.*

### ***Variables independientes excluidas de los modelos tras el análisis previo descriptivo***

- ROE (Return on Equity): porcentaje que mide la rentabilidad que obtienen los accionistas sobre la inversión en la empresa.
- Ingresos: cantidad de dinero que recibe una empresa por los productos que vende o servicios que presta. Medido en millones de dólares (USD).
- Existencia de años intermedios entre año inicial y final (Anos Intermedios): variable categórica que toma el valor 1 si existe un año intermedio de revisión entre el año inicial y final, y el valor 0 si no hay años entre medias. *Fuente: Informes de sostenibilidad de cada una de las empresas (2019-2022).*
- Puntuación Gubernamental de Emisiones Netas Cero (NETZERO SCORE): proporciona una puntuación que mide la robustez de los compromisos de objetivos de cero emisiones netas de los países, basada en el estado de cero emisiones en el proceso legislativo y el año objetivo. La puntuación varía de 0 a 10, siendo 10 el mejor. *Fuente: Bloomberg.*

### **3.4. Modelo**

Un modelo predictivo es una herramienta de análisis cuyo objetivo principal es ofrecer estimaciones e información anticipada basándose en patrones y tendencias que han sido encontradas en datos pasados (Lee et al., 2022). Los modelos predictivos se utilizan para realizar predicciones sobre eventos futuros, mediante el uso de datos históricos (Jain & Raheja, 2021).

Gracias a su capacidad para ofrecer información predictiva y ayudar en la toma de decisiones, este modelo se utiliza en diversos ámbitos, entre ellos el sector financiero y bancario.

Por ello, dada la necesidad de querer estudiar cuáles son las variables independientes que impactan más en la variable dependiente (ROA) se ha decidido hacer un modelo de regresión lineal.

En un modelo de regresión lineal se asume una relación lineal entre la variable dependiente (ROA) y múltiples variables dependientes. De esta forma se pueden observar los distintos factores que afectan al ROA. En R con la función *lm()* se crea el modelo de regresión lineal. En él se especifica

cuál es la variable dependiente que se busca predecir (ROA) y el resto de las variables incluidas en los modelos corresponden a las variables independientes.

En este trabajo de investigación, se han desarrollado un total de cuatro modelos de regresión lineal que permiten analizar qué variables contribuyen de manera significativa al rendimiento del ROA. De esta forma, se identifica si los factores medioambientales impactan al desempeño financiero de las empresas.

En los cuatro modelos, el ROA es la variable dependiente y el resto de las variables numéricas del conjunto de datos normalizados a excepción de los Ingresos y del NETZERO\_SCORE son las variables dependientes. La diferencia es que en cada uno de ellos se incluirá solo una de las puntuaciones (ESG\_SCORE, ENVIRON\_SCORE, GOV\_SCORE o SOCIAL\_SCORE). De esta forma, se evita la alta correlación y redundancia existente entre ellas.

Los cuatro modelos de regresión lineal son:

- Modelo 1: Se incluye el ESG\_SCORE

$$ROA = \beta_1 * EBIT + \beta_2 * EBIT\_ratio + \beta_3 * Objetivo1y2 + \beta_4 * AnoInicio1y2 + \beta_5 * AnoFin1y2 + \beta_6 * Residuos + \beta_7 * UsoEnergiaRenovable + \beta_8 * UsoAgua + \beta_9 * MultasAmbientales + \beta_{10} * CO2\_1y2 + \beta_{11} * GEI\_1y2 + \beta_{12} * CLIMATE\_SCORE + \beta_{13} * UsoElectricidad + \beta_{14} * Derrames + \beta_{15} * CDP\_CLIMATE\_CHANGE\_SCORE + \beta_{16} * SP\_ESG\_RANK * \beta_{17} * ESG\_SCORE$$

- Modelo 2: Se incluye GOV\_SCORE

$$ROA = \beta_1 * EBIT + \beta_2 * EBIT\_ratio + \beta_3 * Objetivo1y2 + \beta_4 * AnoInicio1y2 + \beta_5 * AnoFin1y2 + \beta_6 * Residuos + \beta_7 * UsoEnergiaRenovable + \beta_8 * UsoAgua + \beta_9 * MultasAmbientales + \beta_{10} * CO2\_1y2 + \beta_{11} * GEI\_1y2 + \beta_{12} * CLIMATE\_SCORE + \beta_{13} * UsoElectricidad + \beta_{14} * Derrames + \beta_{15} * CDP\_CLIMATE\_CHANGE\_SCORE + \beta_{16} * SP\_ESG\_RANK * \beta_{17} * GOV\_SCORE$$

- Modelo 3: Se incluye SOCIAL\_SCORE

$$ROA = \beta_1 * EBIT + \beta_2 * EBIT\_ratio + \beta_3 * Objetivo1y2 + \beta_4 * AnoInicio1y2 + \beta_5 * AnoFin1y2 +$$

$$\beta_6 * \text{Residuos} + \beta_7 * \text{UsoEnergiaRenovable} + \beta_8 * \text{UsoAgua} + \beta_9 * \text{MultasAmbientales} +$$

$$\beta_{10} * \text{CO2\_1y2} + \beta_{11} * \text{GEI\_1y2} + \beta_{12} * \text{CLIMATE\_SCORE} + \beta_{13} * \text{UsoElectricidad} + \beta_{14} * \text{Derrames}$$

$$+ \beta_{15} * \text{CDP\_CLIMATE\_CHANGE\_SCORE} + \beta_{16} * \text{SP\_ESG\_RANK} * \beta_{17} * \text{SOCIAL\_SCORE}$$

- Modelo 4: Se incluye ENVIRON\_SCORE

$$\text{ROA} = \beta_1 * \text{EBIT} + \beta_2 * \text{EBIT\_ratio} + \beta_3 * \text{Objetivo1y2} + \beta_4 * \text{AnoInicio1y2} + \beta_5 * \text{AnoFin1y2} +$$

$$\beta_6 * \text{Residuos} + \beta_7 * \text{UsoEnergiaRenovable} + \beta_8 * \text{UsoAgua} + \beta_9 * \text{MultasAmbientales} +$$

$$\beta_{10} * \text{CO2\_1y2} + \beta_{11} * \text{GEI\_1y2} + \beta_{12} * \text{CLIMATE\_SCORE} + \beta_{13} * \text{UsoElectricidad} + \beta_{14} * \text{Derrames}$$

$$+ \beta_{15} * \text{CDP\_CLIMATE\_CHANGE\_SCORE} + \beta_{16} * \text{SP\_ESG\_RANK} * \beta_{17} * \text{ENVIRON\_SCORE}$$

Al ajustar un modelo de regresión lineal en R y utilizar la función de summary() en el objeto resultante del modelo, se obtienen varios estadísticos que se van a analizar para poder interpretar el modelo. Todas las estadísticas proporcionadas, ofrecen una visión completa del modelo ajustado y su capacidad para explicar la variación en los datos (Faraway, 2002). Los resultados incluyen (Montgomery et al., 2012): estimaciones de coeficientes, el error estándar de las estimaciones de coeficientes, el R cuadrado, el R-cuadrado ajustado y el estadístico F.

## 4. RESULTADOS

### 4.1. ANÁLISIS DESCRIPTIVO

El análisis descriptivo facilita la interpretación de patrones y tendencias generales dentro del conjunto de datos. En un principio, la base de datos contiene 56 registros distribuidos en 26 campos distintos, sumando un total de 1.512 observaciones recopiladas en un período de 4 años (2019-2022). Asimismo, se compone de 24 variables numéricas, 1 variable categórica y 1 variable dicotómica. Sin embargo, se ha decidido evaluar el desempeño financiero con el ROA, por lo que, la variable ROE ha sido eliminada desde el inicio ya que son variables muy relacionadas.

Asimismo, es esencial verificar la presencia de valores ausentes en el conjunto de datos, ya que la ausencia de datos puede distorsionar la representatividad de la muestra. Para comprobar si esto ocurría, se ha utilizado el comando de *“is.na()”* en R. Al analizar los resultados, se ha observado la existencia de algún valor perdido. Estos valores han sido sustituidos por las medias de cada una de las variables. Este criterio se ha establecido porque todas ellas son empresas pertenecientes a la industria de las emisiones difíciles de reducir, de ahí que sus variables sean similares.

En la tabla 2 se observan los resultados que se han obtenido en R con el comando *“summary()”*. Con este, se ha realizado un resumen estadístico de cada una de las variables numéricas. Esta nos brinda información de los valores de la media, la mediana, el valor máximo, el valor mínimo, y el primer y tercer cuartil de cada variable.

Tabla 2. Resumen estadístico de la base de datos

Variables	Valor mínimo	1er cuartil	Mediana	Media	3er cuartil	Valor máximo
ROA	-7.86%	2.69%	3.70%	4.79%	5.78%	18.37%
Ingresos	2,322.00	10,282.00	23,610.00	29,044.00	37,690.00	93,328.00
EBIT	-5,387.50	367.10	2,265.40	3,813.00	4,213.60	29,437.00
EBIT_ratio	-15.88%	4.26%	10.44%	11.37%	15.62%	46.36%
Objetivo1y2	20.00%	30.00%	33.00%	34.24%	35.00%	50.00%
AnoInicio1y2	1990	2015	2018	2014	2018	2021
AnoFin1y2	2025	2030	2030	2030	2030	2035
Residuos	31.70	912.90	1,845.00	89,319.80	59,345.00	1,005,650.00
UsoEnergiaRenovable	0.00	988.10	7,479.40	9,405.60	11,521.00	67,202.00
UsoAgua	8,955.00	306,351.00	712,500.00	1,187,636.00	1,235,227.00	6,881,000.00
MultasAmbientales	0.00	3.38	3.50	3.50	3.50	13.00
CO2_1y2	70.69	19,235.85	42,661.35	42,661.35	42,661.35	193,905.00
GEI_1y2	2,313.00	10,792.00	24,850.00	32,032.00	35,022.00	182,384.00
CLIMATE_SCORE	4.17	5.59	5.59	5.59	5.73	7.18
NETZERO_SCORE	2.33	7.67	7.67	7.67	9.00	9.25
UsoElectricidad	2,160.00	4,216.00	13,722.00	19,250.00	28,927.00	60,560.00
Derrames	0.00	5.00	34.55	34.55	34.55	165.00
CDP_CLIMATE_CHANGE_SCORE	2.00	6.00	6.27	6.27	7.00	8.00
SP_ESG_RANK	20.00	53.50	62.46	62.46	76.00	96.00
ESG_SCORE	2.23	4.34	5.03	5.02	5.88	6.69
ENVIRON_SCORE	1.01	4.14	4.92	4.75	5.56	6.19
GOV_SCORE	4.12	5.89	6.29	6.37	7.10	8.74
SOCIAL_SCORE	1.93	2.547	4.16	4.257	5.97	6.95

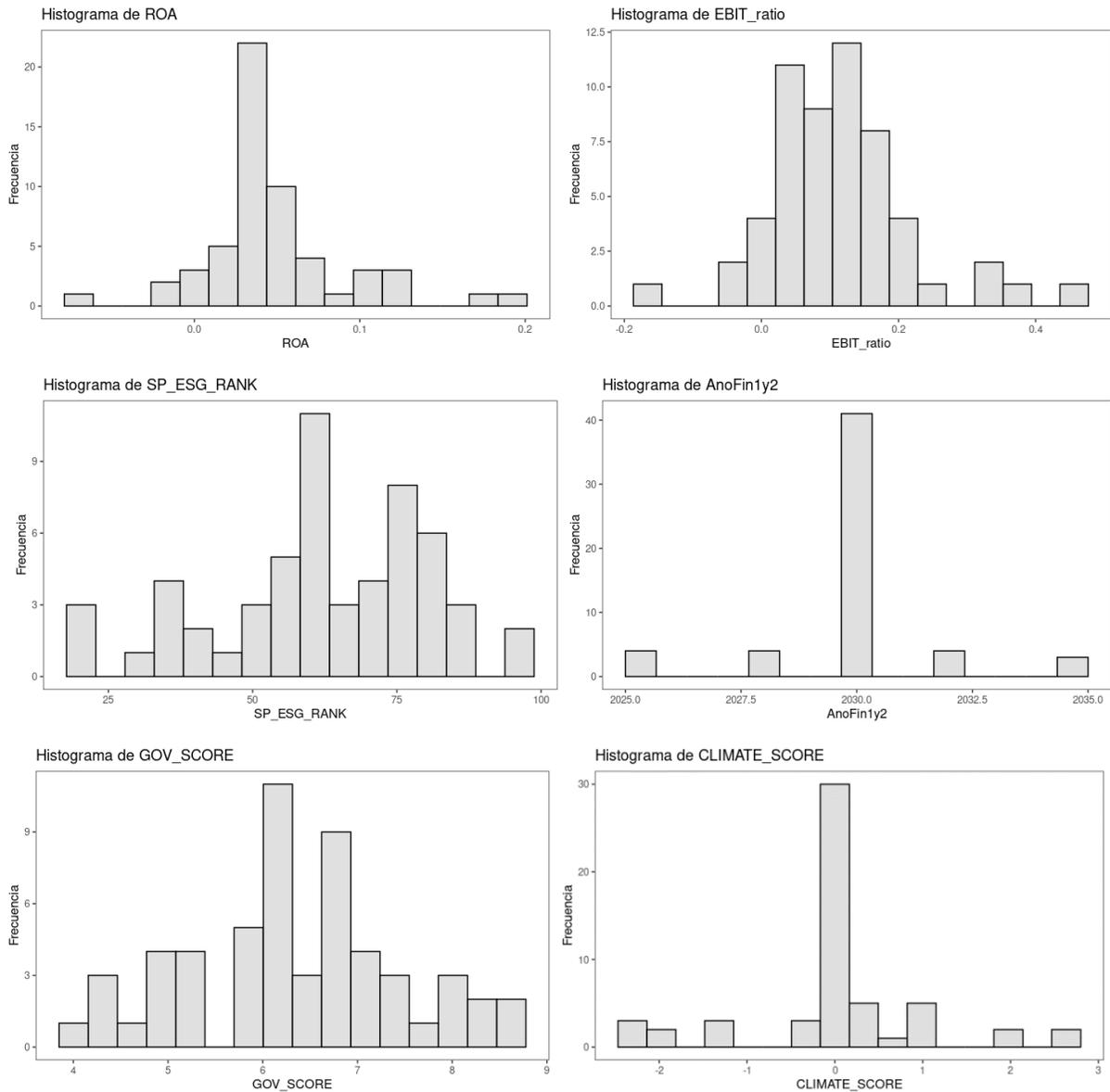
Fuente: Elaboración propia

#### 4.1.1. Distribución de las variables numéricas

En un análisis descriptivo para poder profundizar más sobre el conjunto de datos, es muy importante conocer y profundizar sobre la distribución de las variables. Al final si una variable muestra un sesgo pronunciado o una alta concentración de valores en cierta parte de su rango indica que hay factores o características específicas que influyen y afectan a dicha variable que podrían ser relevantes para el análisis. El sesgo, que también se conoce como asimetría, puede ser tanto positivo como negativo. Una asimetría positiva, o también llamada asimetría a la derecha indica que existen más valores inferiores a la media que superiores. Por el contrario, una asimetría negativa o a la izquierda muestra lo contrario, que hay más valores superiores que inferiores a la media. A continuación, se observa la distribución de cada una de las variables que han sido creadas mediante la función *hist()* de R.

En primer lugar, se observa que las variables del ROA, EBIT Ratio, SP\_ESG\_RANK, AnoFin1y2, GOV\_SCORE y CLIMATE\_SCORE parecen seguir una distribución prácticamente simétrica y normal. Por tanto, se podría afirmar que la media de estas tres variables es representativa para la mayoría de las situaciones y puede servir como referencia para comparar las distintas empresas y comprender el lugar que ocupa cada compañía dentro de la industria.

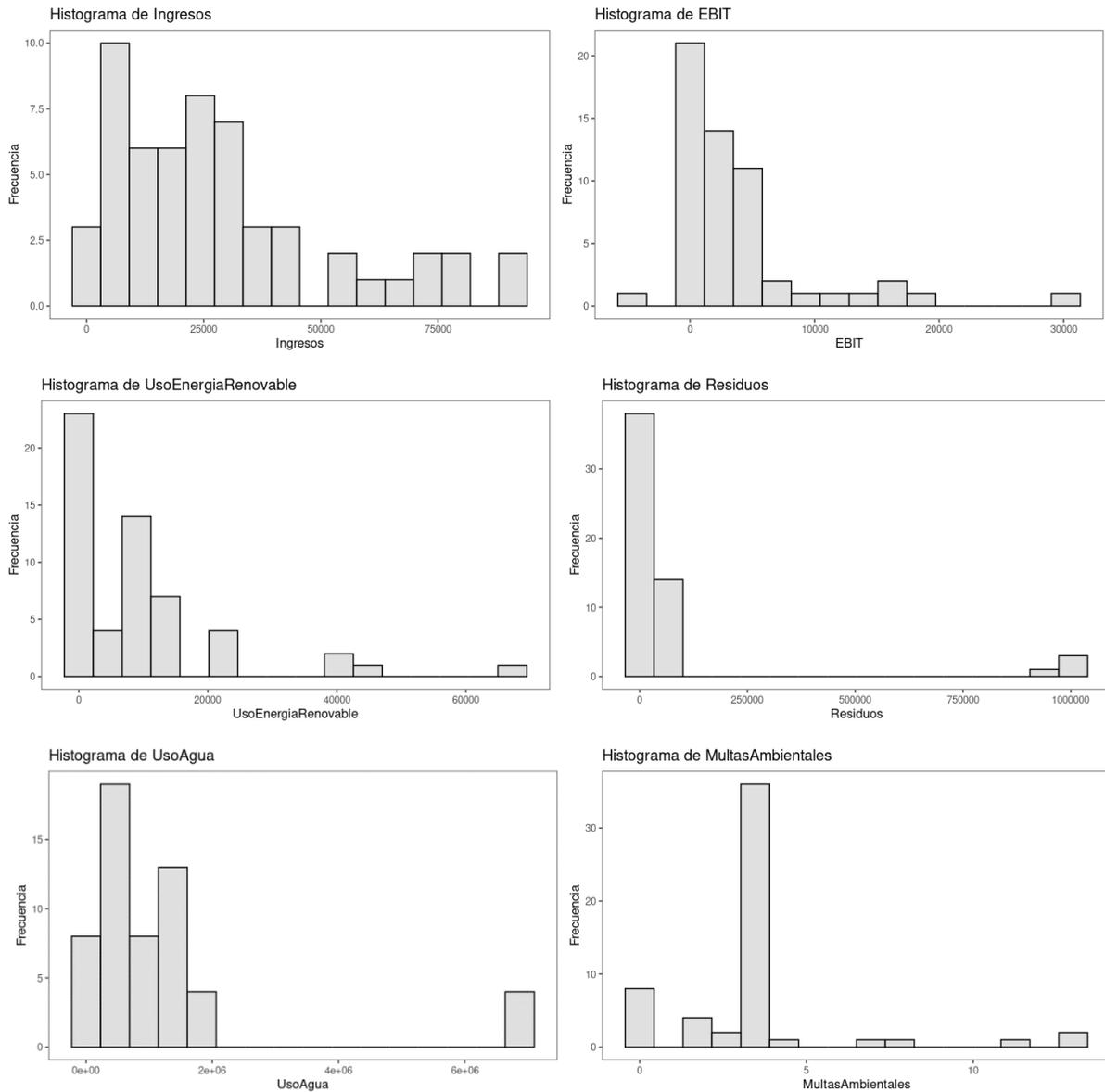
*Figura 4. Variables distribución normal*

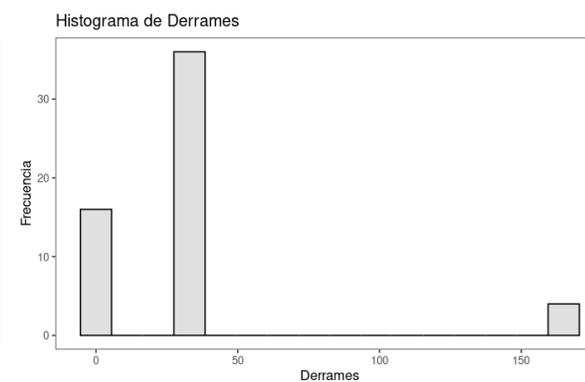
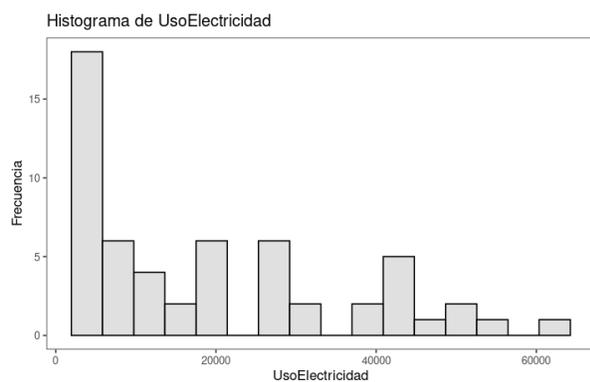
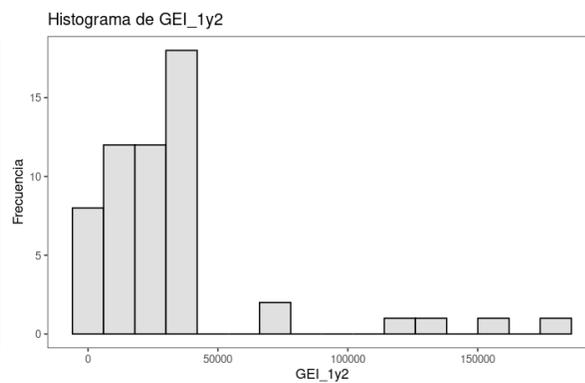
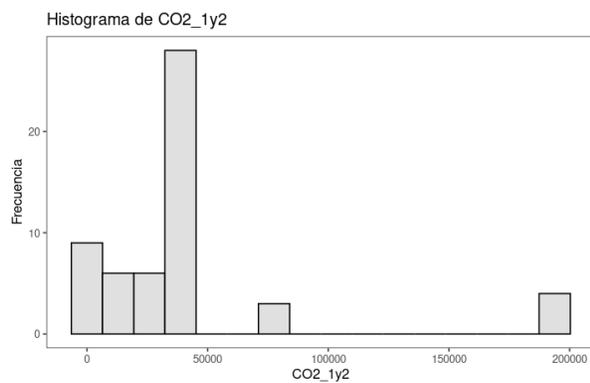


*Fuente: Elaboración propia (RStudio)*

Por otro lado, en la figura 5 se observan las variables que presentan una asimetría a la derecha. Estas son: los ingresos, el EBIT, UsoEnergiaRenovable, Residuos, UsoAgua, MultasAmbientales, CO2\_1y2, GEI\_1y2, UsoElectricidad y Derrames. La presencia de sesgo a la derecha hace que la media no sea un buen representante del valor de los datos, ya que esta influenciado por valores bajos.

*Figura 5. Variables distribución asimétrica a la derecha*

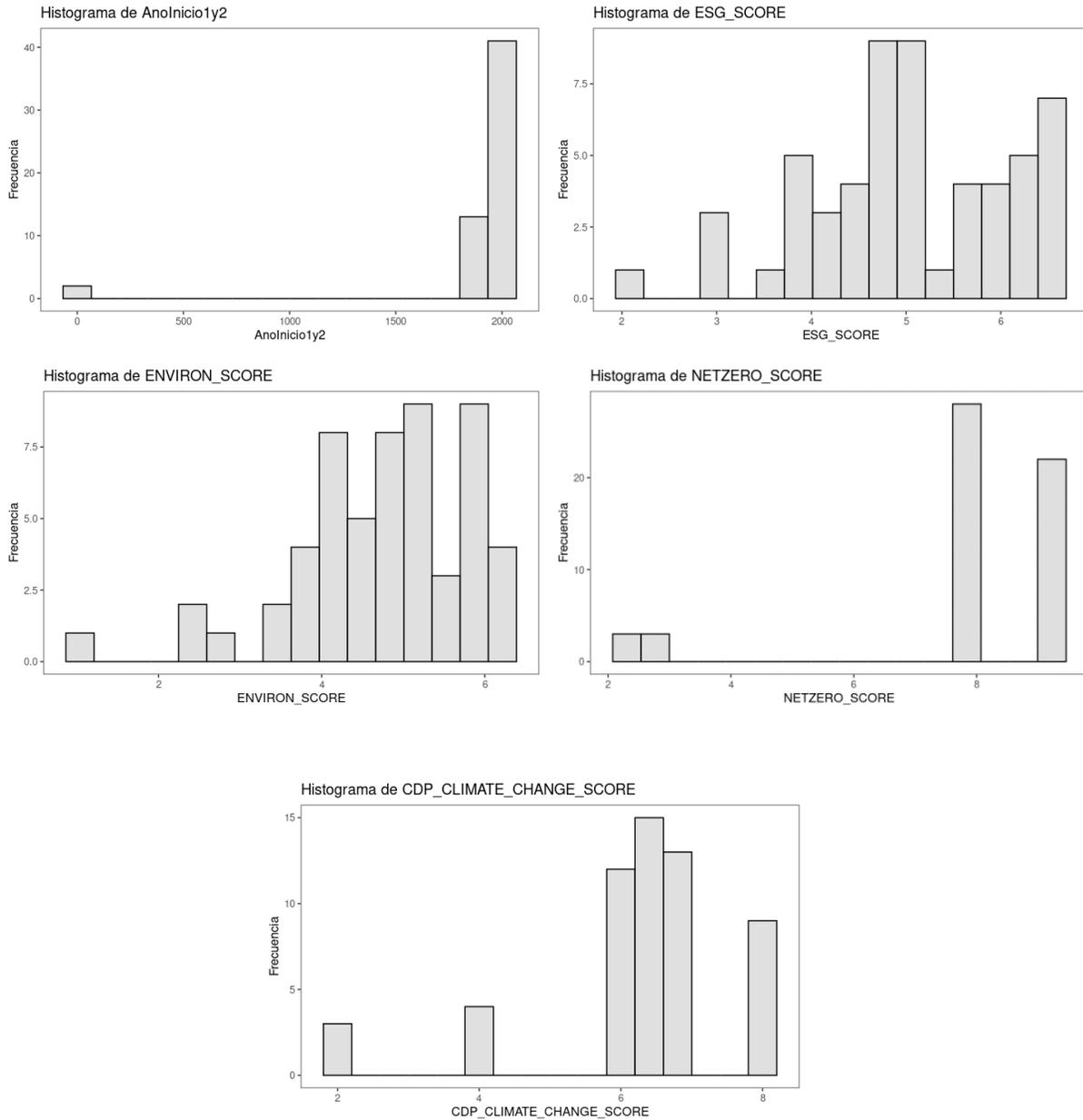




*Fuente: Elaboración propia (RStudio)*

Por el otro lado, aquellas que siguen una distribución asimétrica a la izquierda son AnoInicio\_1y2, ESG\_SCORE, ENVIRON\_SCORE, NETZERO\_SCORE y CDP\_CLIMATE\_CHANGE\_SCORE. En dichas variables, toda la concentración de los valores se encuentra en el extremo izquierdo. Por tanto, las conclusiones extraídas es que las empresas han obtenido una puntuación elevada en estos ámbitos, lo que sugiere que están siguiendo buenas prácticas sostenibles.

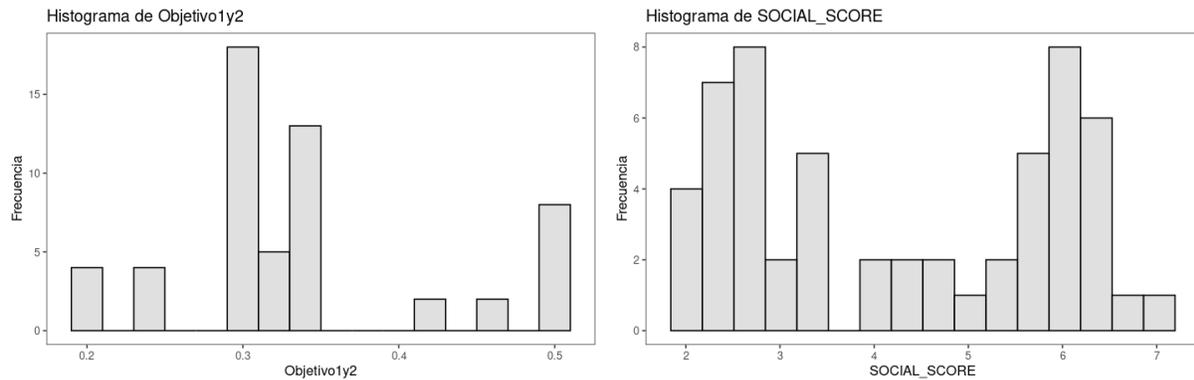
Figura 6. Variables distribución asimétrica a la izquierda



Fuente: Elaboración propia (RStudio)

Finalmente, las variables de Objetivo1y2 y SOCIAL\_SCORE presentan una doble simetría. La distribución se divide prácticamente igual en las dos. Esta bimodalidad que se puede observar en la figura 7, sugiere la existencia de dos categorías distintas en la muestra que están afectando las variables Objetivo1y2 y SOCIAL\_SCORE.

Figura 7. Variable con distribución con doble simetría



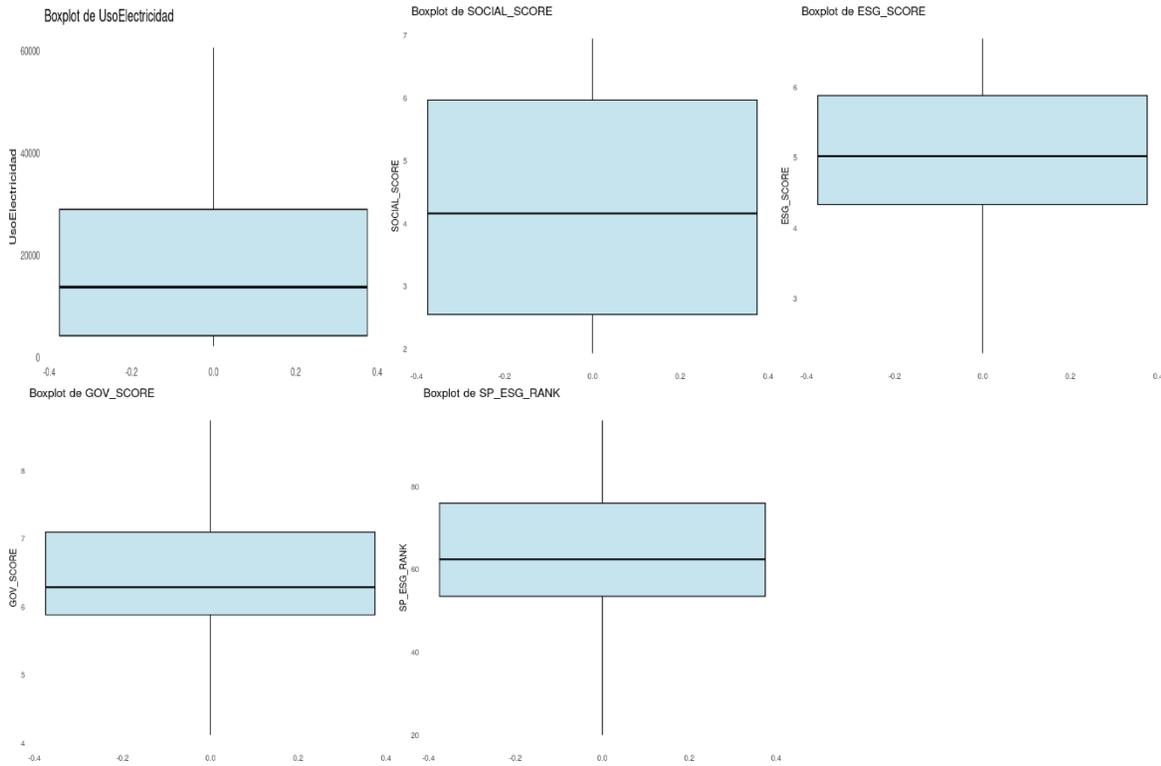
Fuente: Elaboración propia (RStudio)

#### 4.1.2. Valores atípicos

Los valores atípicos o *outliers* son aquellos puntos que difieren de la mayoría de las observaciones en un conjunto de datos. La presencia de estos valores puede afectar al rendimiento y validez de los modelos predictivos distorsionando las conclusiones y disminuyendo la precisión de los resultados obtenidos. No obstante, es importante señalar que un valor atípico no es necesariamente un dato incorrecto o un error, puede simplemente indicar un fenómeno especial e interesante que requiere un análisis más profundo. Por ello, su análisis es fundamental para evaluar si deben de ser excluidos o si son valores que representan casos de interés significativos en ellos mismos.

Las figuras 8 y 9 representan los diagramas de caja de cada una de las variables. Cualquier punto de datos que sitúe fuera de los bigotes, se le considera un *outlier*. En la figura 8, se muestra que las variables de UsoElectricidad, SOCIAL\_SCORE, ESG\_SCORE, GOV\_SCORE y SP\_ESG\_RANK que no tienen valores atípicos.

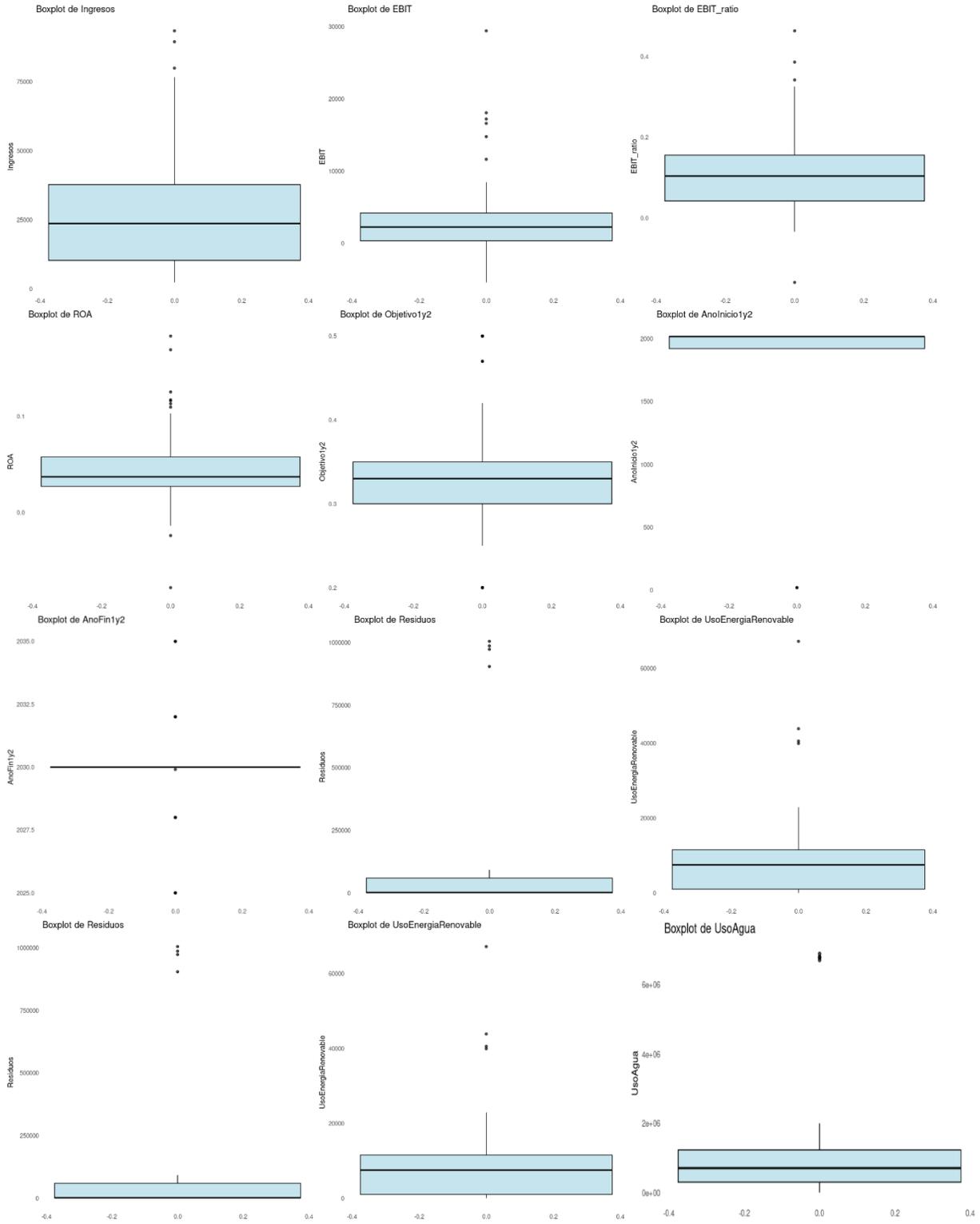
Figura 8. Variables sin valores atípicos

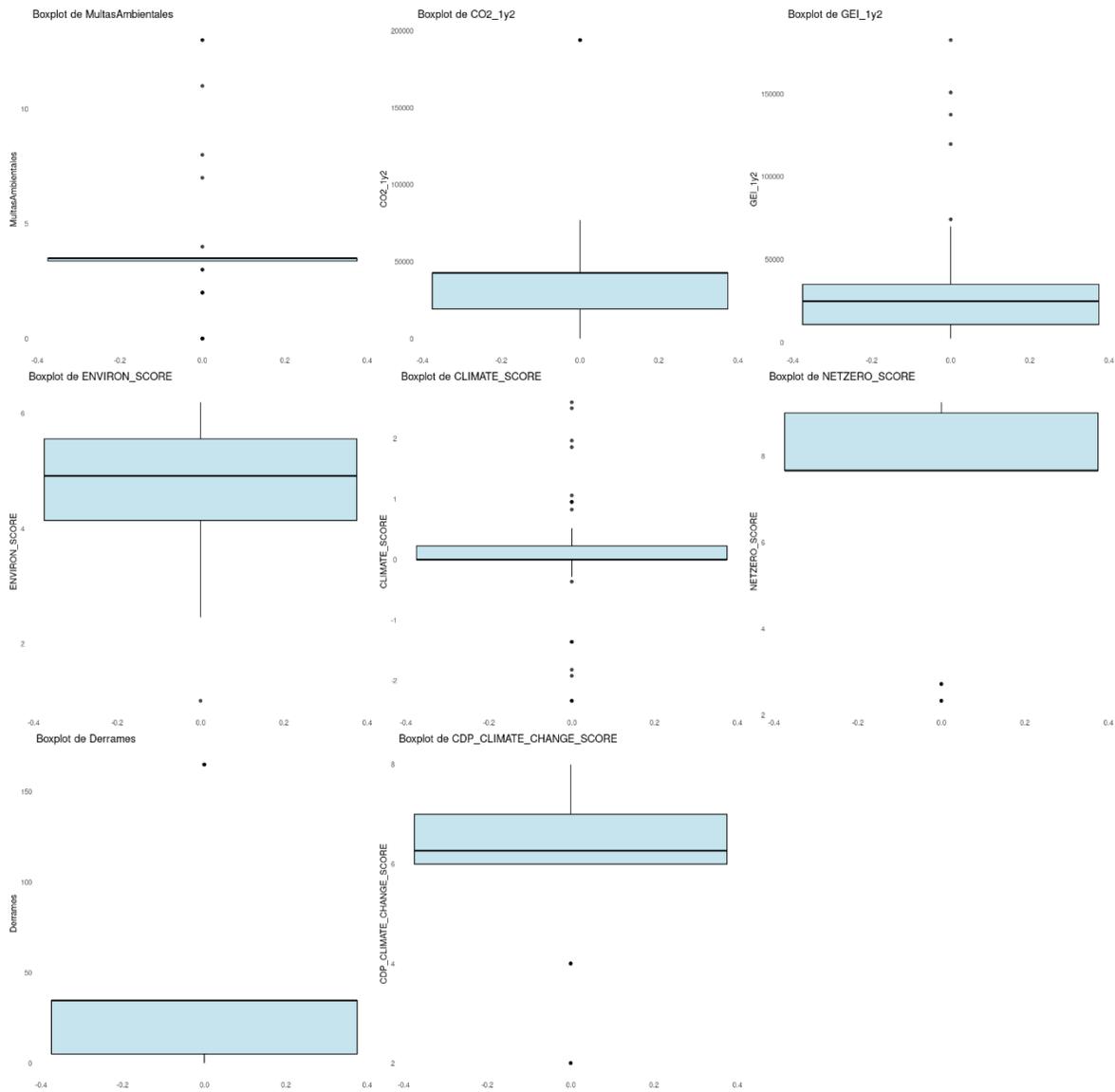


Fuente: Elaboración propia (RStudio)

Por el otro lado, las variables que contienen valores atípicos son: Ingresos, EBIT, EBIT Ratios, ROA, Objetivos\_1y2, AnoInicio1y2, AnoFin1y2, Residuos, UsoEnergiaRenovable, UsoAgua, MultasAmbientales, CO2\_1y2, GEI\_1y2, NETZERO\_SCORE, Derrames, SOCIAL\_SCORE y CDP\_CLIMATE\_CHANGE\_SCORE. En la figura 9, se observan los diagramas de caja de cada una de las variables mencionadas previamente.

Figura 9. Variables con valores atípicos





Fuente: Elaboración propia (RStudio)

Todas estas variables de la figura 9 requieren un análisis más detallado. Para ello, se han recopilados todas las observaciones que se consideran valores atípicos. En la tabla 3, podemos observar el número de valores atípicos que tiene cada una de las variables. Entre todas ellas, suman un total de 128 observaciones las que difieren del resto de valores dentro del conjunto de datos.

*Tabla 3. Número de valores atípicos por variables*

<b>Variables</b>	<b>Número de valores atípicos</b>
<b>ROA</b>	8.00
<b>Ingresos</b>	3.00
<b>EBIT</b>	6.00
<b>EBIT Ratio</b>	4.00
<b>Objetivo1y2</b>	14.00
<b>AnoFin1y2</b>	16.00
<b>Residuos</b>	4.00
<b>UsoEnergiaRenovable</b>	4.00
<b>UsoAgua</b>	4.00
<b>MultasAmbientales</b>	20.00
<b>CO2_1y2</b>	4.00
<b>GEI_1y2</b>	5.00
<b>CLIMATE_SCORE</b>	18.00
<b>NETZERO_SCORE</b>	6.00
<b>Derrames</b>	4.00
<b>CDP_CLIMATE_CHANGE_SCORE</b>	7.00
<b>ENVIRON_SCORE</b>	1.00

*Fuente: Elaboración propia (RStudio)*

Tras haber examinado y analizado los valores atípicos encontrados en el conjunto de datos, hay que determinar si estos deben de ser descartados o si, por el contrario, deben de conservarse. Después de un análisis minucioso, se ha llegado a la conclusión que en esta situación específica los valores atípicos no se van a eliminar. Respecto al total de observaciones de la base de datos, 128 valores no tienen tanto peso como para poder sesgar los resultados de la investigación. Asimismo, son observaciones muy dispersas que demuestran casos muy excepcionales de cada una de las variables afectadas. La base de datos inicial consta de 56 filas, y en 50 de ellas está presente al menos un valor atípico de alguna variable. Por consiguiente, si se eliminan el conjunto de datos se quedaría únicamente con 6 filas, y los modelos de regresión lineal no aportarían mucha información.

### 4.1.3. Normalización de las variables

Con el fin de poder comparar y estudiar las distintas variables de la base de datos, es necesario normalizar los datos. Normalizar implica ajustar los valores medidos en diferentes escalas a una escala común. Esto se puede lograr mediante diferentes técnicas.

Si los datos siguen una distribución normal, se pueden normalizar con la función de R de *scale()*. En este caso se ha aplicado esa normalización para todas aquellas variables que siguen una asimetría entre -0.5 y 0.5. En el caso de que presenten cualquier asimetría sea a la derecha o a la izquierda, se han aplicado distintas técnicas. Para las variables que tienen una asimetría positiva y significativa, es decir a partir de 1, se ha aplicado una transformación logarítmica con la función *log()* en R.

Asimismo, se ha aplicado la técnica de Box-Cox para aquellas variables con asimetría positiva pero no tan significativa, las que se encuentran entre 0.5 y 1. La transformación de Box-Cox emplea un parámetro denominado lambda, que se determina maximizando una función de probabilidad. Esta función ayuda a encontrar el valor de lambda que mejor ajuste los datos a una distribución normal. Finalmente, en la base de datos había variables con una asimetría negativa y significativa. Para aquellas se ha empleado una transformación inversa. La inversión de los datos les permite volverse más simétricos.

En la tabla 4, se puede observar la asimetría que presentaba cada una de las variables y el método de normalización empleado.

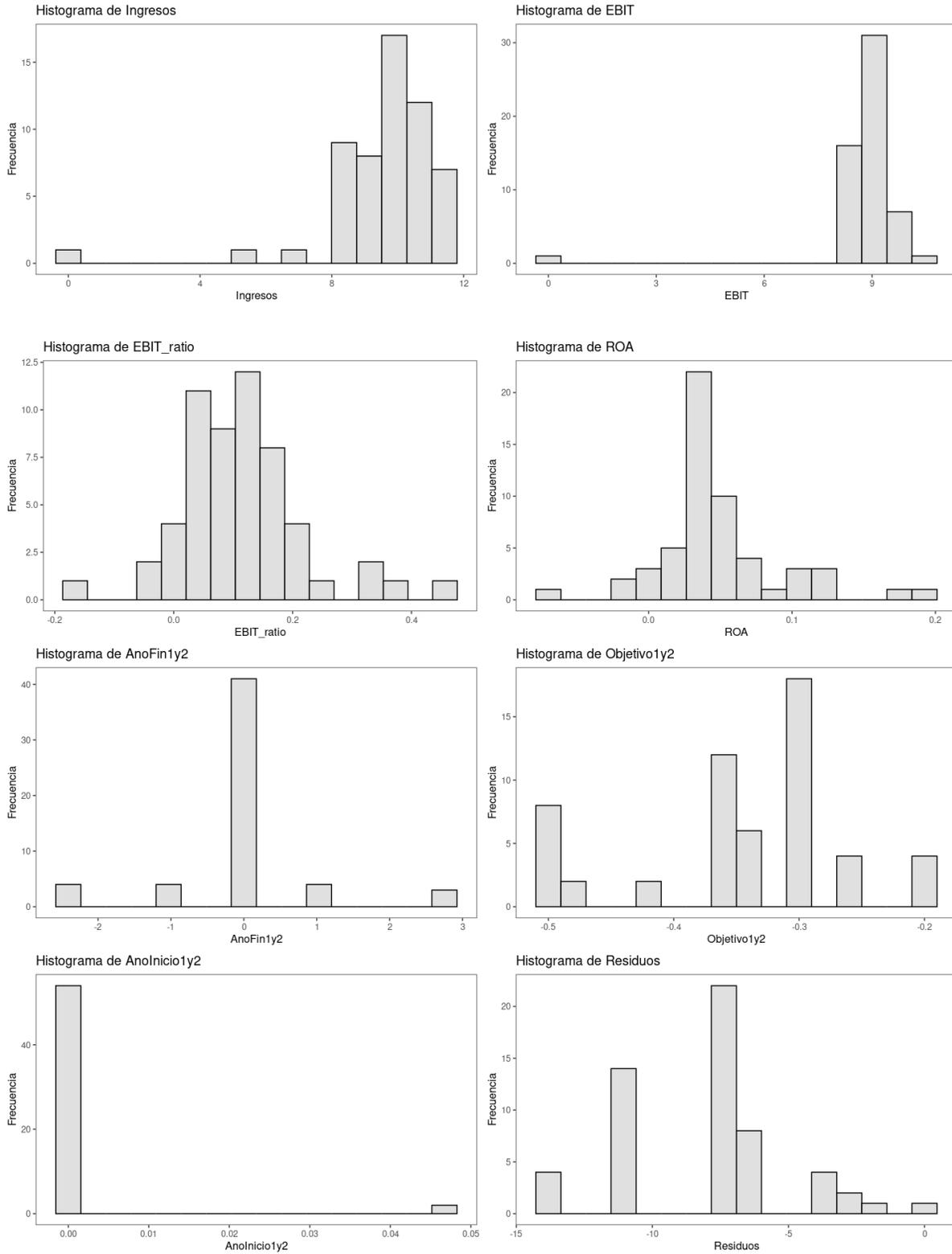
Tabla 4. Asimetría y transformación empleada por variable

<b>Variables</b>	<b>Valor</b>	<b>Transformación aplicada</b>	<b>Razones de asimetría</b>
<b>ROA</b>	0.66	Box-Cox	Asimetría positiva leve
<b>Ingresos</b>	1.10	Logarítmica	Asimetría positiva moderada a fuerte
<b>EBIT</b>	2.31	Logarítmica	Asimetría positiva fuerte
<b>EBIT_ratio</b>	0.88	Box-Cox	Asimetría positiva moderada
<b>Objetivo1y2</b>	0.60	Box-Cox	Asimetría positiva leve
<b>AnoInicio1y2</b>	(4.78)	Inversa	Asimetría negativa muy fuerte
<b>Años_intermedios</b>	0.58	Box-Cox	Asimetría positiva leve
<b>AnoFin1y2</b>	(0.17)	Z-score	Asimetría casi nula
<b>Residuos</b>	3.15	Logarítmica	Asimetría positiva fuerte
<b>UsoEnergiaRenovable</b>	2.40	Logarítmica	Asimetría positiva fuerte
<b>UsoElectricidad</b>	0.73	Box-Cox	Asimetría positiva leve
<b>Derrames</b>	2.44	Logarítmica	Asimetría positiva fuerte
<b>UsoAgua</b>	2.67	Logarítmica	Asimetría positiva fuerte
<b>MultasAmbientales</b>	1.90	Logarítmica	Asimetría positiva fuerte
<b>CO2_1y2</b>	2.32	Logarítmica	Asimetría positiva fuerte
<b>GEI_1y2</b>	2.55	Logarítmica	Asimetría positiva fuerte
<b>CLIMATE_SCORE</b>	(0.12)	Z-score	Asimetría casi nula
<b>NETZERO_SCORE</b>	(1.92)	Inversa	Asimetría negativa fuerte
<b>CDP_CLIMATE_CHANGE_SCORE</b>	(1.44)	Inversa	Asimetría negativa fuerte
<b>SP_ESG_RANK</b>	(0.55)	Inversa	Asimetría negativa leve
<b>ESG_SCORE</b>	(0.40)	Z-score	Asimetría moderada
<b>ENVIRON_SCORE</b>	(1.02)	Inversa	Asimetría negativa moderada
<b>SOCIAL_SCORE</b>	0.04	Z-score	Asimetría casi nula
<b>GOV_SCORE</b>	-0.01	Z-score	Asimetría casi nula

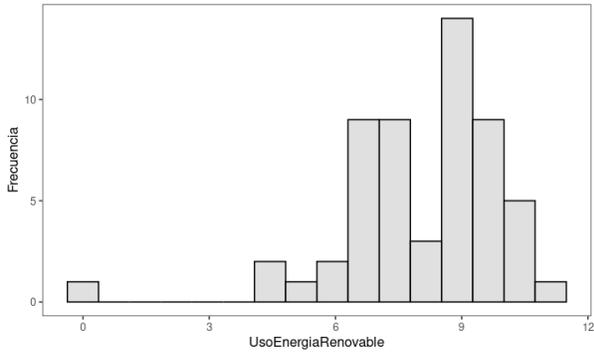
Fuente: Elaboración propia

En la figura 9 se muestran las gráficas de las variables numéricas normalizadas.

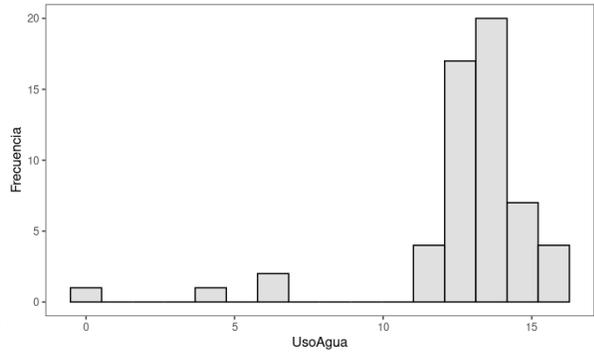
Figura 9. Gráficas de las variables numéricas normalizadas



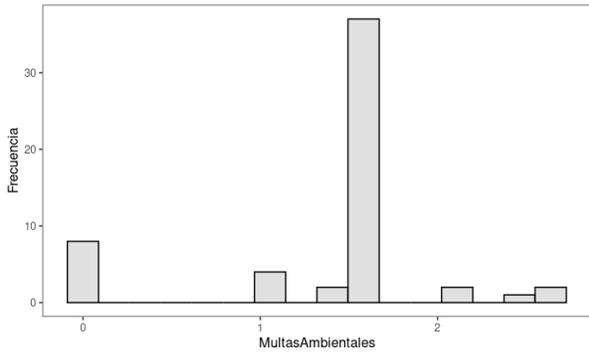
Histograma de UsoEnergiaRenovable



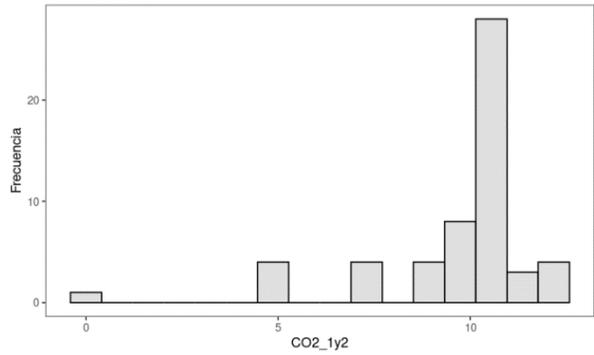
Histograma de UsoAgua



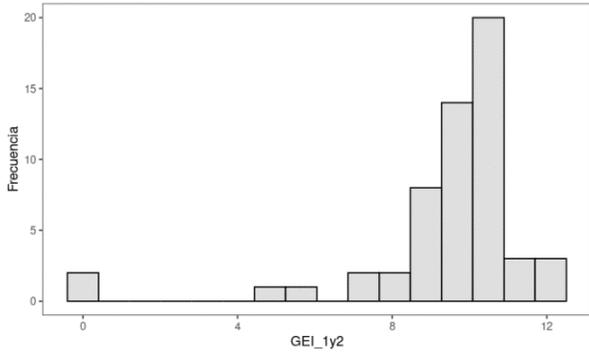
Histograma de MultasAmbientales



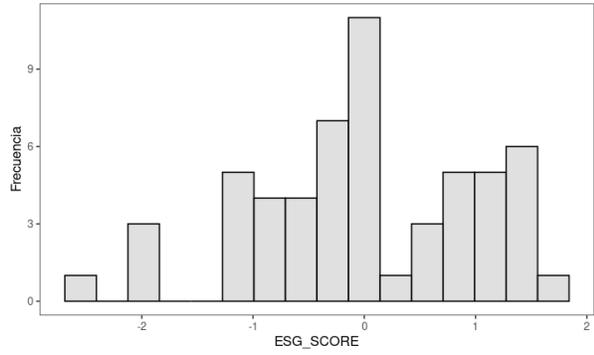
Histograma de CO2\_1y2



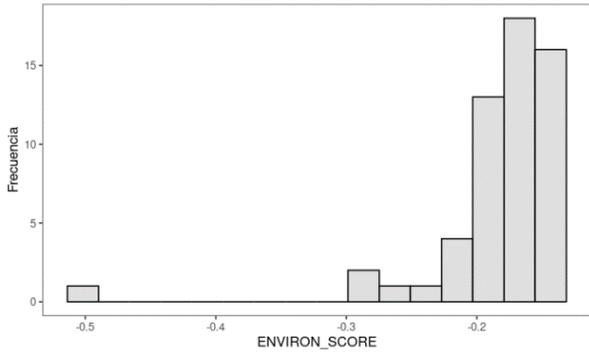
Histograma de GEI\_1y2



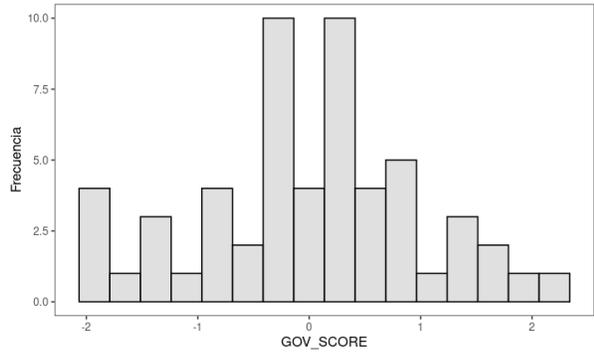
Histograma de ESG\_SCORE

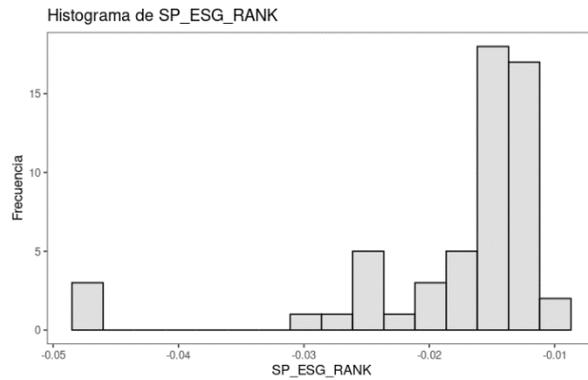
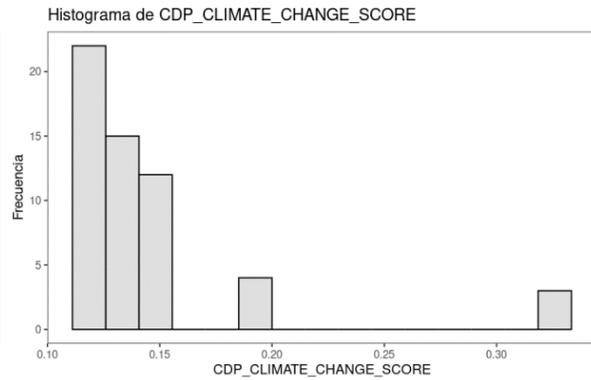
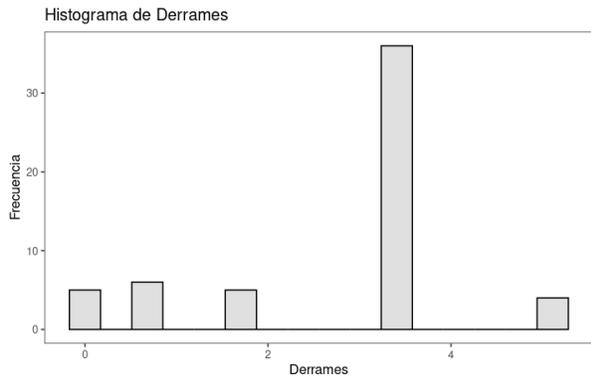
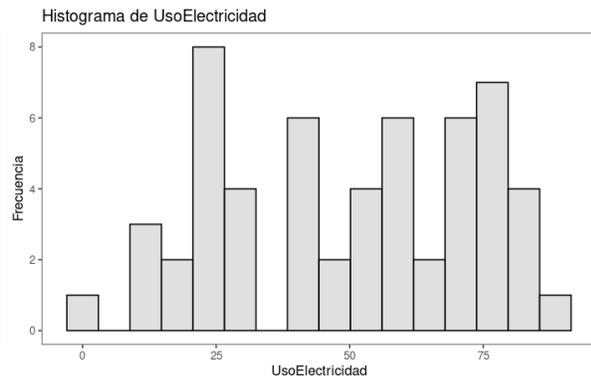
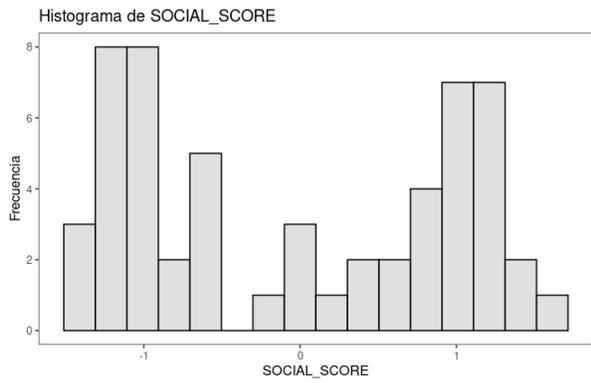
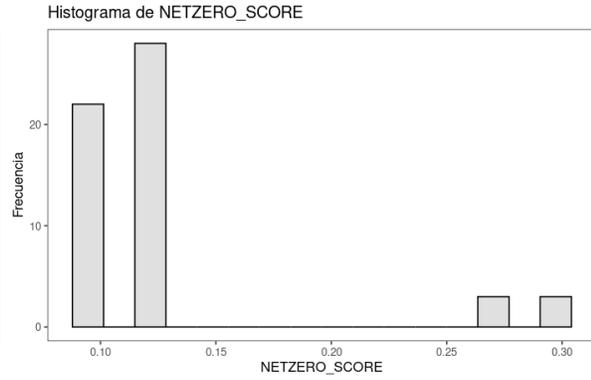
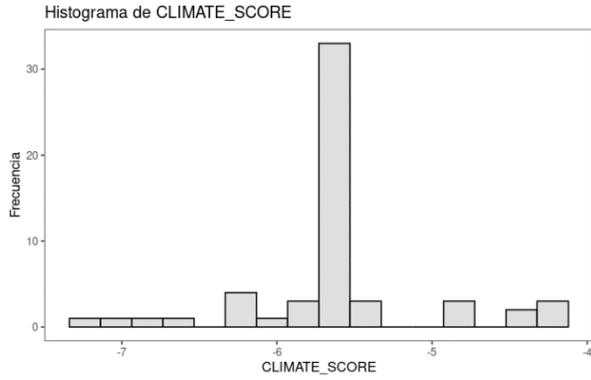


Histograma de ENVIRON\_SCORE



Histograma de GOV\_SCORE





*Fuente: Elaboración propia (RStudio)*

Diferentes conclusiones se extraen de estas distribuciones. En primer lugar, un sesgo a la izquierda en el total de energía renovable empleada (UsoEnergiaRenovable), indica la existencia de desigualdad por parte de las entidades en la adopción de energías renovables. Al estar sesgada a la izquierda, la mayoría de las observaciones se concentran en valores altos lo que implica que muchas entidades están empleando una cantidad alta de energía renovable. Del mismo modo, el consumo de agua (UsoAgua) presenta una asimetría negativa lo que revela que el consumo de agua de las compañías es elevado y se encuentra por encima de la media. Asimismo, en las variables de los alcances 1 y 2 de dióxido de carbono (CO2\_1y2) y de gases de efecto invernadero (GEI\_1y2) también es asimétrica a la izquierda. Esto implica que la mayoría de las entidades tienen emisiones de CO2 y GEI relativamente altas, con algunas entidades significativamente por debajo de esta tendencia.

Los ingresos también presentan una asimetría a la izquierda, esto sugiere la presencia de una desigualdad de los ingresos entre las empresas de la muestra. Lo que revela que hay una pequeña proporción de las empresas de la muestra con unos ingresos muy elevados. Además, un sesgo a la izquierda en las distintas puntuaciones como la puntuación medioambiental (ENVIRON\_SCORE) y la puntuación del rango S&P 500 (SP\_ESG\_RANK), indica que la mayoría de los valores analizados presentan cifras más elevadas. Esto puede sugerir una puntuación más elevada en los distintos conceptos. Por el contrario, las puntuaciones de emisiones netas de CO2 (NETZERO\_SCORE) y las del CDP cambio climático (CDP\_CLIMATE\_CHANGE\_SCORE), tienen una asimetría más a la derecha, lo que indica la presencia de unas puntuaciones más bajas.

En resumen, después de la normalización se observan variables con un sesgo a la izquierda lo que implica una tendencia hacia valores más elevados, y dependiendo de la variable es favorable o no. Mientras que otras, al contrario, tienen una asimetría positiva por lo que son más propensos a valores más bajos. Finalmente, las que siguen una distribución aparentemente normal (EBIT ratio, ROA, porcentaje de los objetivos 1 y 2 (Objetivos\_1y2), año fin de los alcances 1 y 2 (AnoFin1y2), residuos totales (Residuos), número de multas (MultasAmbientales), puntuación ESG (ESG\_SCORE), puntuación gubernamental (GOV\_SCORE), puntuación del riesgo del cambio climático (CLIMATE\_SCORE), puntuación social (SOCIAL\_SCORE), uso de la

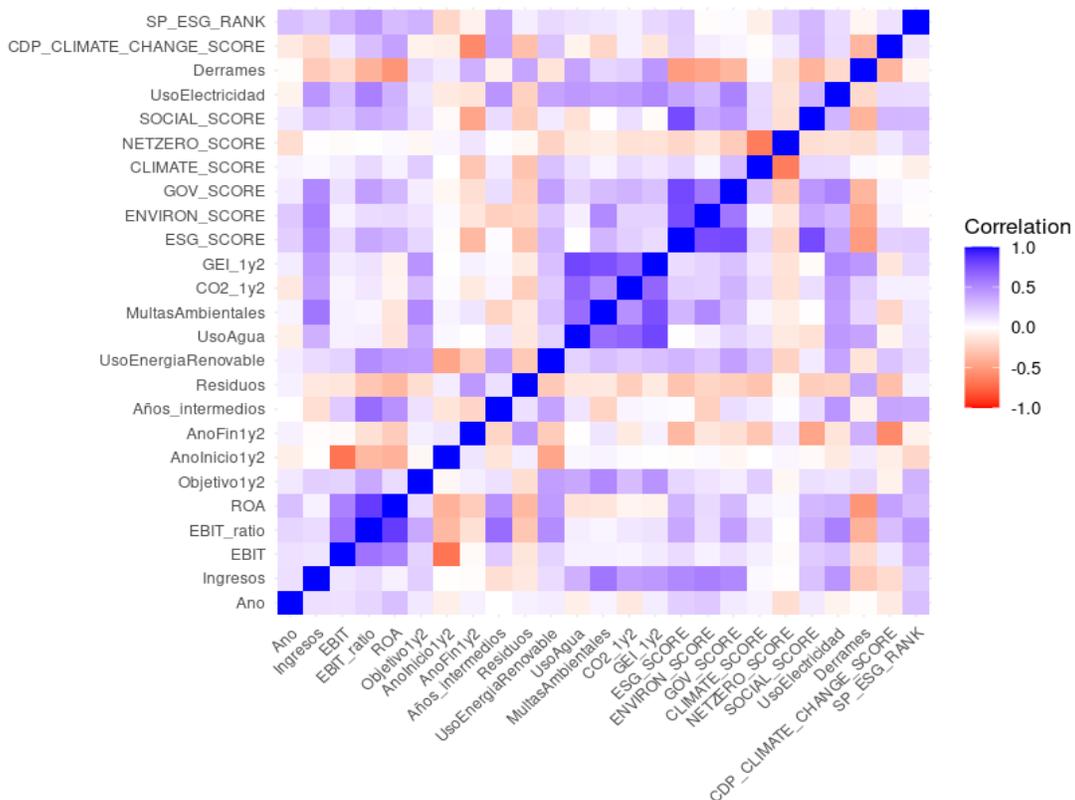
electricidad(UsosElectricidad) y número de derrames (Derrames) sugiere que los datos se distribuyen simétricamente alrededor de la media.

#### 4.1.4. Análisis de correlación

El análisis de correlación es una técnica estadística que permite analizar y cuantificar la relación entre dos variables. Esta técnica permite una comprensión profunda mediante la identificación de patrones y relaciones. Es un análisis que ofrece numerosas ventajas que serán luego muy interesantes para el modelo predictivo. Entre ellas las más destacadas son: la identificación de las variables más relevantes, la predicción de resultados y la validación de hipótesis.

Mediante el comando *cor()* de R se han obtenido los valores de la matriz de correlación y para crear el gráfico que se muestra en la figura 10 se ha empleado comando de *corrplot()*. A través de este se puede ver la intensidad y la dirección de relación entre los pares de variables de la matriz de correlación.

Figura 10. Matriz de correlación



Fuente: Elaboración propia(RStudio)

El gráfico que se ha presentado revela distintas conclusiones respecto a la relación existente entre las variables

Se observa una correlación fuerte positiva entre el EBIT ratio y el ROA (0.84). Esto se debe a que ambos son indicadores financieros que miden aspectos relacionados con la rentabilidad operativa y el uso eficiente de los activos. Una empresa que tiene un alto EBIT es muy probable que su ROA también sea elevado. Por esa misma razón, existe una correlación positiva un poco menor entre el EBIT y el ROA (0.54). La correlación entre el uso de energía renovable (UsoEnergiaRenovable) y el EBIT ratio es altamente positiva (0.49). Esta relación puede ser consecuencia de múltiples factores: empresas que invierten en energía renovable reducen sus costes a largo plazo, lo que lleva a un EBIT más elevado, también empresas con prácticas sostenibles suelen recibir subsidios del gobierno mejorando significativamente su ROA. La relación positiva entre ambas variables se puede asociar a que las prácticas sostenibles mejoran la imagen de la empresa, lo que puede resultar en un incremento en las ventas y por tanto un mayor EBIT ratio. También existe una correlación altamente positiva entre las emisiones de GEI de alcance 1 y 2 (GEI\_1y2), y MultasAmbientales (0.75) y bastante positiva entre las emisiones CO2 de alcance 1 y 2 (CO2\_1y2) MultasAmbientales (0.48). Cuanto mayor sean las emisiones emitidas por las empresas, mayores las penalizaciones obtendrán. De la misma manera, la correlación entre los ingresos y el EBIT ratio es de 0.15, lo que apunta a que mayores ingresos están asociados con un mayor margen operativo.

La variable de la puntuación ESG (ESG\_SCORE) tiene una alta correlación positiva con las puntuaciones de los distintos componentes ya que es una medida agregada derivada de ellos. Sus respectivas correlaciones son: ENVIRON\_SCORE (0.77), GOV\_SCORE (0.79) y SOCIAL\_SCORE (0.78).

Por el contrario, existen variables que tienen una relación inversa. En el caso de la variable número de derrames presenta una correlación negativa moderada con el ROA (-0.54), con la ENVIRON\_SCORE (-0.45), GOV\_SCORE (-0.37), SOCIAL\_SCORE (-0.39) y UsoElectricidad (-0.19). Por un lado, esto nos sugiere que un mayor número de derrames hace que aumenten los costes directos e indirectos asociados con ellos, y esto reduce la eficiencia operativa y la rentabilidad de la empresa resultando en un menor retorno de los activos. Por el otro lado, también

se ve como un mayor número de derrames afecta a la puntuación. Lo que señala que cuántos mayores derrames tengan las compañías, menor es su puntuación. Asimismo, existe correlación negativa entre el EBIT y AnoInicio1y2 (-0.69).

Finalmente, hay variables con colores muy claro, que indican una correlación baja o ninguna. Resalta mucho la variable de emisiones netas (NETZERO\_SCORE) que no tiene apenas correlación con los Ingresos (-0.00839), el EBIT (0.018775), EBIT Ratio (0.005533) y ROA (0.029). Asimismo, las variables AnoInicio1y2, AnoFin1y2 con la variable Objetivo\_1y2 presentan una correlación de 0.03 y -0.073, respectivamente. Lo que sugiere que estas variables no tienen un poder de influencia entre ellas. Sin embargo, no se observa ninguna variable que muestre una correlación mínima con todo el resto de las variables.

La variable objetivo del modelo de regresión lineal es el ROA. A través de este análisis de correlación, se puede observar aquellas variables que presentan una correlación extremadamente baja con el ROA y que, por tanto, no aportan valor. Los resultados de la matriz de correlaciones muestran una relación muy cercana a 0, apenas perceptible entre el ROA y los ingresos (0.064) y el ROA y NETZERO\_SCORE (0.029).

## **4.2. MODELO REGRESIÓN LINEAL**

### **4.2.1. Análisis de la multicolinealidad**

La multicolinealidad hace referencia a la existencia de una relación de dependencia lineal fuerte entre dos o más variables independientes en un modelo de regresión lineal. Es una parte esencial del diagnóstico de un modelo de regresión lineal, ya que un alta multicolinealidad puede afectar a la fiabilidad e interpretación de los coeficientes estimados.

Para su diagnóstico, se ha aplicado la función del código de R de *vif()* al modelo de regresión lineal completo inicial. El modelo de regresión lineal completo tiene el ROA como variable dependiente y el resto de las variables numéricas del conjunto de datos normalizado como independientes. Se han excluido las variables de emisiones netas (NETZERO\_SCORE) y de los ingresos por tener una correlación cercana a cero con el ROA. El VIF representa el Factor de Inflación de la Varianza

y sirve para cuantificar la multicolinealidad en un modelo de regresión. Un valor VIF superior a 10 indica una correlación muy elevada lo que sugiere problemas graves de multicolinealidad.

Los resultados VIF obtenidos para cada una de las variables en el modelo de este trabajo de investigación son los presentados en la tabla 5:

*Tabla 5. Resultados VIF del modelo normalizado*

<b>Variables</b>	<b>VIF</b>
<b>EBIT</b>	5.00
<b>EBIT_ratio</b>	6.32
<b>Objetivo1y2</b>	5.66
<b>AnoInicio1y2</b>	4.79
<b>AnoFin1y2</b>	3.27
<b>Residuos</b>	1.99
<b>UsoEnergiaRenovable</b>	5.68
<b>UsoElectricidad</b>	8.26
<b>Derrames</b>	5.68
<b>UsoAgua</b>	4.91
<b>MultasAmbientales</b>	7.90
<b>CO2_1y2</b>	2.81
<b>GEI_1y2</b>	7.85
<b>CLIMATE_SCORE</b>	1.79
<b>CDP_CLIMATE_CHANGE_SCORE</b>	2.34
<b>SP_ESG_RANK</b>	2.13
<b>ESG_SCORE</b>	53.21
<b>ENVIRON_SCORE</b>	7.65
<b>SOCIAL_SCORE</b>	19.48
<b>GOV_SCORE</b>	11.97

*Fuente: Elaboración propia (RStudio)*

Como se observa en los resultados, existen variables que presentan unos valores VIF significativamente altos superiores a 10, lo que indica una posible multicolinealidad. Estas son: ESG\_SCORE (53.21), SOCIAL\_SCORE (19.48), GOV\_SCORE (11.97) y ENVIRON\_SCORE (7.65) que no llega a 10. La razón de ello es que el ESG SCORE es una medida agregada de las puntuaciones medioambientales, sociales y gubernamentales. Por tanto, al introducir tanto la

puntuación agregada como las de cada uno de los componentes individuales se introduce una alta redundancia. Asimismo, entre ellas son variables muy correlacionadas que evalúan diferentes aspectos de la sostenibilidad.

La solución para evitar que la elevada multicolinealidad distorsione la interpretación del modelo y reduzca la precisión va a ser la realización de cuatro modelos de regresión lineal. En los cuatro modelos, el ROA es la variable dependiente y el resto de las variables numéricas del conjunto de datos normalizados a excepción de los Ingresos y del NETZERO\_SCORE son las variables dependientes. La diferencia es que en cada uno de ellos se incluirá solo una de las puntuaciones (ESG\_SCORE, ENVIRON\_SCORE, GOV\_SCORE o SOCIAL\_SCORE). De esta forma, se evita la alta correlación y redundancia existente entre ellas.

#### **4.2.2. Resultados obtenidos en cada modelo de regresión lineal**

Tras la formulación de las hipótesis y el análisis descriptivo de los datos, se han desarrollado un total de cuatro modelos de regresión lineal con la función  $lm()$  que permiten analizar qué variables contribuyen de manera significativa al rendimiento del ROA.

A continuación, se van a presentar, analizar e interpretar los resultados que se han obtenido de los modelos que han sido previamente definidos. Para evaluar el nivel de significación de cada una de las variables y observar estadísticos útiles que facilitan su interpretación, se ha empleado la función *summary()*.

La función *summary()* del código R devuelve estadísticos que permiten analizar los resultados de los modelos elaborados con una mayor profundidad. El intercepto y los coeficientes de cada una de las variables independientes contienen información sobre la relación lineal entre la variable dependiente (ROA) y variables independientes. Por un lado, el intercepto, representa el valor estimado de la variable dependiente (ROA) cuando todas las variables independientes son iguales a cero, es decir, cuando no hay efecto de las variables independientes en la variable objeto de estudio. Por otro, los coeficientes, indican cómo cambia la variable dependiente en respuesta a un cambio unitario en la variable independiente correspondiente, manteniendo el resto de las variables independientes constantes. Un coeficiente positivo significa que un aumento en el valor de la

variable independiente está asociado con un incremento en el valor de la variable dependiente y viceversa. Es importante tener en cuenta que los coeficientes se interpretan en relación con las unidades de las variables involucradas. Asimismo, se observan los niveles de significación de cada una de las variables independientes. Todas aquellas que tienen un p-valor menor al 5.00% son significativas y permiten rechazar la hipótesis nula.

Toda esta información está recogida en la tabla 6 y presenta los valores de cada modelo para cada una de las variables que están en la ecuación.

Tabla 6. Resultados modelos regresión lineal

	Modelo 1	Modelo 2	Modelo 3	Modelo 4
<i>(Intercept)</i>	0.0421 (0.071)	0.0448 (0.0708)	0.0465 (0.0721)	0.0496 (0.0728)
<b>EBIT</b>	0.0040 (0.0047)	0.0040 (0.0047)	0.0039 (0.0048)	0.0038 (0.0048)
<b>EBIT_ratio</b>	0.3610 (0.0609)**	0.3700 (0.0614)**	0.3543 (0.0619)**	0.3570 (0.0615)**
<b>Objetivo1y2</b>	0.0962 (0.0567)	0.1132 (0.0589)	0.0854 (0.0609)	0.0944 (0.0582)*
<b>AnoInicio1y2</b>	-0.2872 (0.6740)	-0.3345 (0.6737)	0.2048 (0.6758)	-0.1994 (0.6757)
<b>AnoFin1y2</b>	-0.0021 (0.0042)	-0.0009 (0.0040)	-0.0016 (0.0046)	-0.0009 (0.0041)
<b>Residuos</b>	-0.0007 (0.0012)	-0.0006 (0.0012)	-0.0009 (0.0012)	-0.0008 (0.0012)
<b>UsoEnergiaRenovable</b>	0.0024 (0.0017)	0.0040 (0.0029)	0.0029 (0.0030)	0.0031 (0.0029)
<b>UsoAgua</b>	-0.0021 (0.0020)	-0.0015 (0.0019)	-0.0018 (0.0022)	-0.0015 (0.0019)
<b>MultasAmbientales</b>	-0.0057 (0.0094)	-0.0079 (0.0090)	-0.0084 (0.0091)	-0.0071 (0.0106)
<b>CO2_1y2</b>	-0.0011 (0.0019)	-0.0010 (0.0018)	-0.0013 (0.0019)	-0.0014 (0.0019)
<b>GEI_1y2</b>	-0.0075 (0.0032)*	-0.0079 (0.0032)*	-0.0070 (0.0032)*	-0.0071 (0.0032)*
<b>CLIMATE_SCORE</b>	0.0046 (0.0058)	0.0036 (0.0059)	0.0052 (0.0058)	0.0051 (0.0059)
<b>UsoElectricidad</b>	-0.0006 (0.0002)*	-0.0006 (0.0002)*	-0.0005 (0.0002)*	-0.0006 (0.0002)*
<b>Derrames</b>	-0.0086 (0.0038)*	-0.0092 (0.0039)*	-0.0072 (0.0036)*	-0.0076 (0.0041)
<b>CDP_CLIMATE_CHANGE_Si</b>	0.0548 (0.0816)	0.0404 (0.0829)	0.0625 (0.0821)	0.0643 (0.0821)
<b>SP_ESG_RANK</b>	0.4012 (0.4103)	0.5124 (0.4057)	0.4257 (0.4259)	0.4712 (0.4130)
<b>ESG_SCORE</b>	0.0043 (0.0044)			
<b>GOV_SCORE</b>		0.0051 (0.0043)*		
<b>SOCIAL_SCORE</b>			0.0014 (0.0045)	
<b>ENVIRON_SCORE</b>				0.0195 (0.0846)
Número de observaciones	1,008	1,008	1,008	1,008
Error estándar residual	0.02028	0.02016	0.01947	0.02052
R-cuadrado	0.8522	0.8539	0.8279	0.8487
R-cuadrado ajustado	0.7861	0.7886	0.8028	0.781
Estadístico F	12.89	13.07	32.99	12.54
P-valor	6.58E-11	5.33E-11	2.93E-16	1.00E-10

Errores estándar entre paréntesis. \* y \*\*, denotan significación al 5% y al 1%, respectivamente.

Fuente: Elaboración propia (RStudio)

Según los resultados de la tabla 6, se concluye que en todos los modelos las variables del EBIT Ratio, la emisión de los gases de efecto invernadero (GEI\_1y2), y el uso de la electricidad tienen un p-valor menor al 5%, lo que implica una influencia significativa en el ROA. Según el signo que presenten sus coeficientes, sugiere que un aumento en cualquiera de estas variables está asociado a un incremento o una disminución en el ROA. En el caso del EBIT ratio, este refleja mejoras en la rentabilidad operativa y la eficiencia en el uso de los activos, por lo que tiene un impacto positivo en el ROA. Por otro lado, las empresas con altas emisiones enfrentan multas y mayores costes por inversiones en tecnologías limpias, reduciendo su rentabilidad y el ROA. Los aumentos en los costes de electricidad incrementan los costes operativos, reduciendo el margen operativo y el ingreso neto, provocando una disminución en el ROA.

Respecto a estos resultados, se puede concluir que la emisión de gases de efecto invernadero (GEI\_1y2) y el uso de la electricidad (UsoElectricidad) son significativas en todos los modelos e implican que la hipótesis 1 y 2 sean respaldadas por los resultados obtenidos en todos los modelos. De ahí, que se pueda afirmar que una reducción de emisiones y una gestión eficiente del consumo eléctrico experimentan mejoras en el desempeño financiero. Este estudio confirma la tesis de Porter (1991) que afirmaba que las empresas que reducen sus emisiones en la atmósfera son las que mayor eficiencia productiva. Y “La Teoría de Recursos y Capacidades” en la que se examinó cómo las inversiones medioambientales proactivas mejoran la competitividad empresarial

En los modelos 1, 2 y 3 se observa una relación negativa y significativa a un nivel de significación del 5% entre el ROA y el número de derrames (Derrames). Su impacto en ROA se relaciona con que un aumento en el número de derrames incrementa los costes operativos y de limpieza y por consecuencia disminuye el ROA.

Asimismo, en el modelo 2 se observa una relación significativa a un nivel de significación del 5% y positiva entre el ROA y la puntuación gubernamental (GOV\_SCORE). Una puntuación gubernamental más elevada indica que las empresas tienden a ser más transparentes y responsables y eso fortalece la confianza de los inversores y partes interesadas, mejorando el desempeño financiero. Por tanto, la hipótesis 4 es también respaldada por el modelo 2, que sugiere que empresas con una mayor puntuación en gobernanza tienen un mejor desempeño financiero. Este

resultado es consistente con previos estudios. Un informe de Deloitte indicó que las empresas con mayores resultados ESG tienen menores costes de capital, mejorando su rentabilidad en general.

Por su parte, en el modelo 4 existe una relación positiva y significativa para un nivel de significación del 5% entre el ROA y los objetivos establecidos de alcance 1 y 2 (Objetivo\_1y2). Establecer un mayor porcentaje objetivo de reducción de emisiones mejora la eficiencia operativa, financiera y la reputación, incrementando así el ROA.

Finalmente, después de completar la investigación se ha rechazado la hipótesis 3. No se presentan evidencias significativas que empleo de energías renovables mejoren el desempeño del ROA. El traslado a energías renovables a menudo se justifica desde una perspectiva de sostenibilidad a largo plazo, incluyendo beneficios ambientales, sociales y de gobernanza (ESG). Sin embargo, todos estos beneficios no se reflejan inmediatamente en los indicadores financiero como el ROA, que suele enfocarse en el desempeño a corto y medio plazo.

Al observar la tabla 6, es relevante ver que los resultados presentados por el modelo 4 son muy similares a los del modelo 3. La diferencia principal entre ellos es que en el modelo 3 se incluía únicamente la puntuación social (SOCIAL\_SCORE), mientras que en el 4 solo la puntuación medioambiental (ENVIRON\_SCORE). El resto de las variables eran iguales. Esto indica que ni la puntuación social (SOCIAL\_SCORE) ni la medioambiental (ENVIRON\_SCORE) han aportado una mejor significativa al modelo final, de ahí sus resultados similares.

En términos generales, los cuatro modelos son robustos y esto queda evidenciado bajo los distintos resultados obtenidos de cada uno de ellos. Los cuatro modelos presentan un error residual bajo, lo que sugiere que apenas existen diferencias entre los valores observados y los ajustados, indicando una precisión. Los valores del R-cuadrado y R-cuadrado ajustado evidencian un buen ajuste. En todos ellos, el modelo explica más del 75% de la variabilidad del ROA. Finalmente, los resultados de unos estadísticos F elevados y p-valores extremadamente bajos indican la significancia global del modelo en el ROA.

En resumen, tras haber analizado los resultados de cada modelo se han analizado cuáles son las variables medioambientales que tienen un mayor impacto en el rendimiento financiero de las empresas pertenecientes a las industrias con emisiones difíciles de reducir. En la tabla 7, se presenta un resumen de cuáles son las variables incluidas en cada modelo y las hipótesis respaldadas por cada uno de ellos.

Tabla 7. Resumen de las variables e hipótesis respaldadas por los modelos de regresión lineal

Modelos	VARIABLES INCLUIDAS EN LOS MODELOS NORMALIZADOS	HIPÓTESIS RESPALDADAS
<b>Modelo 1</b>	EBIT EBIT_ratio Objetivo1y2 AnoInicio 1y2 AnoFin 1y2 Residuo UsoEnergiaRenovable UsoAgua MultasAmbientales CO2_1y2 GEI_1y2 CLIMATE_SCORE UsoElectricidad Derrames CDP_CLIMATE_CHANGE_SCORE SP_ESG_RANK ESG_SCORE	H1,H2
<b>Modelo 2</b>	EBIT EBIT_ratio Objetivo1y2 AnoInicio 1y2 AnoFin 1y2 Residuo UsoEnergiaRenovable UsoAgua MultasAmbientales CO2_1y2 GEI_1y2 CLIMATE_SCORE UsoElectricidad Derrames CDP_CLIMATE_CHANGE_SCORE SP_ESG_RANK GOV_SCORE	H1,H2,H4
<b>Modelo 3</b>	EBIT EBIT_ratio Objetivo1y2 AnoInicio 1y2 AnoFin 1y2 Residuo UsoEnergiaRenovable UsoAgua MultasAmbientales CO2_1y2 GEI_1y2 CLIMATE_SCORE UsoElectricidad Derrames CDP_CLIMATE_CHANGE_SCORE SP_ESG_RANK SOCIAL_SCORE	H1,H2
<b>Modelo 4</b>	EBIT EBIT_ratio Objetivo1y2 AnoInicio 1y2 AnoFin 1y2 Residuo UsoEnergiaRenovable UsoAgua MultasAmbientales CO2_1y2 GEI_1y2 CLIMATE_SCORE UsoElectricidad Derrames CDP_CLIMATE_CHANGE_SCORE SP_ESG_RANK ENVIRON_SCORE	H1,H2

Fuente: Elaboración propia (Rstudio)

## 5. DISCUSIÓN, CONCLUSIONES Y LIMITACIONES

Este trabajo de investigación se ha centrado en analizar el impacto de las prácticas de gestión medioambiental en el desempeño financiero, mediante el ROA, de empresas europeas pertenecientes a sectores con emisiones difíciles de reducir. Los hallazgos de este estudio amplían la literatura existente relacionada con las prácticas ambientales y el desempeño financiero. Se ofrece una perspectiva cuantitativa enfocada únicamente en industrias de emisiones difíciles de reducir: aluminio, acero, química y cemento.

A continuación, se van a presentar las principales conclusiones obtenidas que proporcionan una guía clara sobre las variables críticas que pueden tener un impacto significativo en el desempeño financiero de las empresas altamente contaminantes.

En primer lugar, los modelos analizados han respaldado la hipótesis de que una reducción de las emisiones de gases de efecto invernadero mejora el desempeño financiero. Esto coincide con “La Hipótesis de Porter”, que sugiere que la eficiencia productiva y la reducción de emisiones están relacionadas (Burnett & Hansen, 2008; Fride et al., 2015). Una menor cantidad de emisiones no mejora únicamente la sostenibilidad ambiental, sino que también contribuye a la eficiencia operativa y a la reducción de costes, mejorando la rentabilidad de las empresas.

Asimismo, un mayor consumo de electricidad se asocia negativamente con el ROA. Por tanto, la eficiencia en el uso de la energía es crucial para la rentabilidad corporativa. Finalmente, una alta puntuación en gobernanza se correlaciona positivamente con el ROA. Unas prácticas de gobernanza robustas y transparentes mejoran la confianza de los inversores y accionistas, lo que puede conducir a un mejor rendimiento financiero.

Los resultados de este estudio tienen importantes implicaciones para las empresas pertenecientes a industrias con emisiones difíciles de reducir. Con todo ello, se recomienda a las empresas pertenecientes a estos sectores que inviertan en tecnologías que les ayuden a reducir las emisiones GEI, optimicen el consumo energético para mejorar la eficiencia operativa y fortalezcan la gobernanza corporativa. La transición hacia la sostenibilidad requiere inversiones significativas y un compromiso a largo plazo. Se ha estimado que serían necesarias inversiones de varios billones de dólares cada año para financiar la acción climática y lograr los Objetivos de Desarrollo

Sostenible (ODS) de la ONU para 2030 (Barua, 2020). Otra estimación indica que se requerirían entre 300 y 500 mil millones de dólares al año entre 2030 y 2050 para establecer una infraestructura resiliente al clima (Adhikari & Chalkasra, 2021).

Una forma de ayudar y conseguir cumplir con dichas recomendaciones es mediante la financiación de transición. La financiación de transición se refiere a una amplia categoría de actividades, productos, servicios (incluida la gestión de riesgos financieros relacionados con el clima y medio ambiente), instrumentos y mecanismos en el ámbito financiero que están vinculados a inversiones en actividades empresariales e industriales que pueden crear un impacto sostenible y positivo en la sociedad y el medio ambiente, incluyendo la tierra, el agua, los recursos de biodiversidad, el aire y las personas (Sachs et al., 2019).

Sin embargo, para acceder a una financiación de transición, las empresas deben de mostrar un compromiso sólido con la sostenibilidad, y proporcionar reportes detallados sobre su desempeño en termino ESG. Asimismo, deben de cumplir con las normativas ambientales vigentes y en ocasiones se les puede exigir algunas certificaciones específicas como ISO 14001. Además, las empresas deben de manifestar un historial financiero sólido y un plan de negocios que muestre cómo la financiación de transición va a contribuir a la viabilidad económica del proyecto. Por su parte, los bancos evalúan tanto la sostenibilidad como la viabilidad financiera de los proyectos, asegurando que los fondos se utilicen de manera efectiva para promover la sostenibilidad y mitigar los riesgos financieros y ambientales.

Investigaciones futuras podrían estudiar los desafíos que presentan las pequeñas y medianas empresas (PYMES) para conseguir la financiación necesaria para la transición a prácticas sostenibles. La financiación hacia la sostenibilidad para las pequeñas y medianas empresas (PYMES) es crucial para cumplir con los objetivos climáticos. Las PYMES representan alrededor del 40% de las emisiones de gases de efecto invernadero en el sector empresarial en los países de la UE y aproximadamente el 50% a nivel global (ITC, 2021; OCDE, 2023). Sin embargo, la mayoría de estas empresas se encuentran en etapas iniciales de su camino a lograr emisiones netas y enfrentan dificultades significativas para acceder a la financiación necesaria para esta transición.

La limitada financiación es uno de los principales obstáculos de las PYMEs en su acción climática. Muchas instituciones financieras requieren datos detallados sobre el desempeño de sostenibilidad de las compañías para gestionar riesgos y cumplir con los requisitos. Las PYMES, a menudo, no tienen las capacidades necesarias para proporcionar estos datos de manera efectiva. Por otro lado, la falta de información y de conciencia sobre la transición al neto cero por parte de las PYMES limita la demanda de financiación sostenibles. Muchas de ellas no ven el caso de negocio para las inversiones sostenibles debido a los altos costes iniciales y los retornos inciertos a largo plazo.

Por ello, si se quiere alcanzar la transición hacia emisiones netas cero de gases de efecto invernadero (GEI) para el 2050, es necesario que tanto el sector público como las instituciones financieras contribuyan a crear un ecosistema propicio que facilite la inversión de las PYMEs en sostenibilidad.

El análisis de esta investigación se enfrenta a limitaciones que han podido influir en la precisión y alcance de los resultados y conclusiones obtenidas.

En primer lugar, la selección de la muestra empleada para el estudio. Esta se limita a 14 empresas de sectores con emisiones difíciles de reducir. Aunque es representativa de las industrias seleccionadas, puede no ser generalizable a todas las empresas europeas de este sector. Por tanto, las conclusiones obtenidas podrían no ser aplicables a un conjunto más amplio de empresas con características y contextos diferentes.

En segundo lugar, es una muestra compuesta únicamente por empresas europeas, lo cual puede introducir sesgos geográficos. Las políticas regulatorias, los incentivos gubernamentales y las condiciones económicas varían considerablemente entre las regiones, lo que puede influir en los resultados y limitar la extrapolación de los hallazgos a nivel mundial.

Asimismo, el horizonte temporal del estudio abarca solo el periodo de 2019 a 2022. Los cambios en las políticas ambientales, tecnológicas y económicas más allá de este período podrían influir significativamente en los resultados. Así eventos externos como son la pandemia del COVID-19

o la guerra de Rusia y Ucrania han podido haber distorsionado los resultados financieros y prácticas de sostenibilidad.

Por otro lado, la calidad y disponibilidad de los datos pueden haber influido en los resultados del estudio. La precisión y la actualización de la información se puede ver limitada al ser datos obtenidos de informes de sostenibilidad y bases de datos financieros. Además, la falta de datos homogéneos y consistentes para todas las empresas y periodos considerados pueden perturbar a la robustez de las conclusiones. También, las empresas suelen presentar los datos de manera más favorable a su imagen pública lo que puede generar un sesgo de información.

Finalmente, las métricas y métodos utilizados para medir el desempeño ambiental y financiero pueden variar entre las empresas y estudios, lo que dificulta las comparaciones directas y consistencia en los resultados.

Todas estas limitaciones subrayan la necesidad de que futuras investigaciones aborden estos desafíos para obtener una comprensión más completa y precisa del impacto de las prácticas de gestión ambiental en el desempeño financiero de los sectores de emisiones difíciles de reducir.

A pesar de ello, este trabajo proporciona una base sólida que permite entender y profundizar sobre el impacto de las prácticas sostenibles en el desempeño financiero de empresas de sectores de emisiones difíciles de reducir. Asimismo, sugiere áreas de mejora para que futuras investigaciones puedan abordar estos obstáculos y proporcionar un análisis más completo.

## 6. ANEXO

En este anexo se muestra el código de R que se ha utilizado para la creación de los modelos de regresión lineal utilizados en el trabajo. Para ello se han seguido varios pasos:

- Paso 1: Instalación de los paquetes y librerías*
- Paso 2: Importar la base de datos*
- Paso 3: Eliminación de los valores desconocidos (NA)*
- Paso 4: Elaboración del resumen estadístico*
- Paso 5: Análisis de los valores atípicos*
- Paso 6: Detección de la asimetría de las variables*
- Paso 7: Normalización de las variables*
- Paso 8: Matriz de correlación*
- Paso 9: Análisis de multicolinealidad*
- Paso 10: Creación de los modelos de regresión lineal*
- Paso 11: Resultados de los modelos de regresión lineal*

A continuación, se presenta más a detalle cada uno de los pasos, junto con los códigos implementados.

### **Paso 1: Instalación de los paquetes y librerías que se van a utilizar a lo largo de todo el trabajo**

```
install.packages("readxl")
install.packages("dplyr")
install.packages("tidyr")
install.packages("car")
install.packages("e1071")
install.packages("ggplot2")
```

```
library(dplyr)
library(tidyr)
library(e1071)
library(car)
library(ggplot)
library(readxl)
```

### **Paso 2: Importar la base de datos y transformar las variables según su naturaleza (numérica, texto..) y renombrarlas**

- Importación de la base de datos y transformación

```
datosbuenos <- read_excel("datosbuenos.xlsx",
col_types = c("numeric", "text", "numeric", "numeric", "numeric", "numeric", "numeric",
"numeric", "text", "text", "numeric", "numeric", "numeric", "numeric", "numeric", "numeric",
```

```
"numeric", "numeric", "numeric", "numeric", "numeric", "numeric", "numeric", "numeric",  
"numeric", "numeric", "numeric", "numeric", "numeric"))
```

```
View(datosbuenos)
```

```
datos<-datosbuenos
```

- Renombrar las variables

```
datos <- datos %>%  
  rename(%_Objetivo_Scope1_2 =Objetivo1y2 , Ano_Inicio_1_2 = AnoInicio1y2,  
  Ano_Final_1_2 = AnoFin1y2, TOTAL_WASTE=Residuos,  
  RENEW_ENERGY_USE=UsoEnergiaRenovable ,  
  NUM_ENVIRON_FINES=MultasAmbientales, TOTAL_WATER_USE=UsoAgua,  
  CO2_SCOPE_1_2_LOCATION_BASED=CO2_1y2,  
  GHG_SCOPE_1_2_LOCATION_BASED=GEI_1y2,  
  ENVIRONMENTAL_SCORE=ENVIRON_SCORE, GOVERNANCE_SCORE=GOV_SCORE,  
  GOVT_CLIMATE_RISK_SCORE=CLIMATE_SCORE,  
  GOVT_NET_ZERO_SCORE=NETZERO_SCORE, ELECTRICITY_USED=UsoElectricidad,  
  NUMBER_SPILLS=Derrames, SP_R_ESG_RANK=SP_ESG_RANK )
```

### **Paso 3: Eliminar los valores NA y sustituirlos por la media**

- Se excluye la variable del ROE que ya no va a ser utilizada

```
datos <- subset(datos, select = -ROE)
```

- Se cuenta el número de valores NA

```
na_count <- sum(is.na(datos))
```

```
if(na_count >0){cat("El data frame contiene",na_count,"valores NA.\n")}else{cat("El data frame  
no contiene valores NA.\n")}
```

- Se sustituyen por la media

```
datos <-datos %>%mutate(across(where(is.numeric),~replace_na(.,mean(.,na.rm =TRUE))))
```

### **Paso 4: Realización de un resumen estadístico**

```
summary(datos)
```

### **Paso 5: Obtener el número de valores atípicos**

- Calcular el número total de valores atípicos

```
indices_atipicos <- function(x) { bp_stats <- boxplot.stats(x) atipicos <- which(x %in%
bp_stats$out) return(atipicos) }
```

```
atipicos_por_columna <- datos %>% summarise(across(where(is.numeric), ~
list(indices_atipicos(.))))
```

```
total_atipicos <- sum(sapply(atipicos_por_columna, function(x) length(unlist(x)))) cat("Número
total de valores atípicos en el dataframe:", total_atipicos, "\n")
```

- Mostrar las observaciones atípicas por variables

```
for (col in names(atipicos_por_columna)) { atipicos <- unlist(atipicos_por_columna[[col]])
cat("Observaciones atípicas en la variable", col, "son:", atipicos, "\n") }
```

- Realización de boxplot para observar los valores atípicos:

```
for (variable in names(datosnorm)) { if (is.numeric(datosnorm[[variable]])) {
```

```
# Abrir una nueva ventana gráfica
```

```
dev.new()
```

```
p <- ggplot(datosnorm, aes_string(y = variable)) + geom_boxplot(fill = "lightblue", color =
"black") + ggtitle(paste("Boxplot de", variable)) + theme_minimal() + theme( plot.title =
element_text(hjust = 0.5), axis.title.x = element_blank(), axis.title.y = element_text(size = 12),
panel.grid.major = element_blank(), panel.grid.minor = element_blank() ) print(p)
readline(prompt = "Presione [Enter] para continuar con el siguiente boxplot") }
```

```
indices_atipicos <- function(x) { bp_stats <- boxplot.stats(x) atipicos <- which(x %in%
bp_stats$out) return(atipicos) } # Aplicar la función a todas las columnas numéricas del
dataframe y almacenar los resultados atipicos_por_columna <- datosnorm %>%
summarise(across(where(is.numeric), ~ list(indices_atipicos(.))))
```

## Paso 6: Detectar la asimetría de las variables

- Realización de los histogramas antes de normalizar:

```
for (variable in names(datos)) {
if (is.numeric(datos[[variable]])) {
dev.new() # Abrir una nueva ventana gráfica
p <- ggplot(datos, aes_string(x = variable)) +
geom_histogram(binwidth = diff(range(datos[[variable]])) / 15, fill = "lightgrey", color =
"black", alpha = 0.7) +
labs(title = paste("Histograma de", variable), x = variable, y = "Frecuencia") +
```

```
theme_bw() + # Tema con fondo blanco
theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) # Sin líneas de
cuadrícula
```

```
print(p)
readline(prompt = "Presiona [Enter] para ver el siguiente histograma")
}
}
```

- Detectar la cantidad de asimetría de cada una de las variables:

```
datos_numericos <- datos[,sapply(datos,is.numeric)]
asimetria <- apply(datos_numericos,2,skewness)
print(asimetria)
```

### **Paso 7: Normalizar las variables según el tipo adecuada**

```
datosnorm <- datos
```

- Normalización SCALE:

```
datosnorm$AnoFin1y2 <- scale(datosnorm$AnoFin1y2)
datosnorm$GOV_SCORE <- scale(datosnorm$GOV_SCORE)
datosnorm$SOCIAL_SCORE <- scale(datosnorm$SOCIAL_SCORE)
datosnorm$CLIMATE_SCORE <- scale(datosnorm$CLIMATE_SCORE)
datosnorm$ESG_SCORE <- scale(datosnorm$ESG_SCORE)
```

- Normalización LOGARÍTMICA:

```
log_transform <- function(x){# Añadir una constante para evitar log(0) o log(valores
negativos) constante <- -1 -min(x,na.rm =TRUE)x_ajustado <- x +constante
return(log(x_ajustado))}
```

```
datosnorm <- datosnorm %>%
mutate(
Ingresos = log_transform(Ingresos),
MultasAmbientales = log_transform(MultasAmbientales),
Residuos = log_transform(Residuos),
CO2_1y2= log_transform(CO2_1y2),
EBIT= log_transform(EBIT),
UsoEnergiaRenovable= log_transform(UsoEnergiaRenovable),
GEI_1y2 = log_transform(GEI_1y2),
TOTAL_WATER_USE= log_transform(TOTAL_WATER_USE),
Derrames= log_transform(Derrames)
)
```

- Normalización BOX-PLOT:

## BOX-COX

```
box_cox_transform <-function(x){# Ajuste para valores no positivos (Box-Cox requiere valores
positivos)constante <-1 -min(x,na.rm =TRUE)x_ajustado <-x +constante # Encontrar el mejor
lambda para la transformaciónbc <-boxcox(x_ajustado ~1,lambda =seq(-2,2,0.1))lambda_optimo
<-bc$x[which.max(bc$y)]# Aplicar la transformación Box-Cox con el lambda
óptimoif(lambda_optimo ==0){return(log(x_ajustado))}else{return((x_ajustado^lambda_optimo
-1)/lambda_optimo)}
```

```
datosnorm <- datosnorm %>%
mutate(
ROA = box_cox_transform(ROA),
Años_intermedios= box_cox_transform(Años_intermedios),
Objetivo1y2 = box_cox_transform(Objetivo1y2),
UsoElectricidad = box_cox_transform(UsoElectricidad),
EBIT_ratio= box_cox_transform(EBIT_ratio)
)
```

- Normalización INVERSA:

```
inverse_transform <-function(x){return(1/(x +1))}
datosnorm<-datosnorm
%>%mutate(ENVIRON_SCORE=inverse_transform(ENVIRON_SCORE),SP_ESG_RANK=in
verse_transform(SP_ESG_RANK),AnoInicio1y2
=inverse_transform(AnoInicio1y2),NETZERO_SCORE
=inverse_transform(NETZERO_SCORE),CDP_CLIMATE_CHANGE_SCORE=inverse_transf
orm(CDP_CLIMATE_CHANGE_SCORE))
```

- Realización de los histogramas después de normalizar:

```
for (variable in names(datosnorm)) {
if (is.numeric(datosnorm[[variable]])) {
dev.new() # Abrir una nueva ventana gráfica
p <- ggplot(datosnorm, aes_string(x = variable)) +
geom_histogram(binwidth = diff(range(datosnorm[[variable]])) / 30, fill = "lightgrey", color =
"black", alpha = 0.7) +
labs(title = paste("Histograma de", variable), x = variable, y = "Frecuencia") +
theme_bw() + # Tema con fondo blanco
theme(panel.grid.major = element_blank(), panel.grid.minor = element_blank()) # Sin líneas de
cuadrícula

print(p)
```

```

readline(prompt = "Presiona [Enter] para ver el siguiente histograma")
}
}

```

### **Paso 8: Realizar la matriz de correlación**

```

datosnormnum <- datosnorm[sapply(datosnorm, is.numeric)]
correlation_matrix <- cor(datosnormnum, use = "complete.obs")
melted_correlation_matrix <- melt(correlation_matrix)

ggplot(data = melted_correlation_matrix, aes(x = Var1, y = Var2, fill = value)) +
  geom_tile() +
  scale_fill_gradient2(low = "red", high = "blue", mid = "white", midpoint = 0, limit = c(-1, 1),
space = "Lab", name = "Correlation") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 45, vjust = 1, size = 10, hjust = 1),
axis.text.y = element_text(size = 10)) +
  coord_fixed()

```

### **Paso 9: Detectar multicolinealidad**

- Crear el modelo completo con los datos normalizados

```

modelonormalizado <- lm(ROA ~ . - Ano - Empresa - Años_intermedios - Ingresos -
NETZERO_SCORE , data = datosnorm)

```

- Aplicar la función VIF

```

vif_values <-vif(modelonormalizado)

print vif_values)

```

### **Paso 10: Crear los 4 modelos de regresión lineal**

- Modelo 1:

```

modelonormalizado1 <- lm(ROA ~ . - Ano - Empresa - Años_intermedios - Ingresos -
NETZERO_SCORE - GOV_SCORE -SOCIAL_SCORE -ENVIRON_SCORE, data =
datosnorm)

```

- Modelo 2:

```

modelonormalizado2 <- lm(ROA ~ . - Ano - Empresa - Años_intermedios - Ingresos -
NETZERO_SCORE ESG_SCORE -SOCIAL_SCORE -ENVIRON_SCORE, data = datosnorm)

```

- Modelo 3:

```
modelonormalizado3 <- lm(ROA ~ . - Ano - Empresa - Años_intermedios - ESG_SCORE -  
GOV_SCORE - ENVIRON_SCORE, data = datosnorm)
```

- Modelo 4:

```
modelonormalizado4 <- lm(ROA ~ . - Ano - Empresa - Años_intermedios - ESG_SCORE -  
GOV_SCORE - SOCIAL_SCORE, data = datosnorm)
```

### **Paso 11: Resultados de los distintos modelos de regresión lineal**

```
summary(modelonormalizado1)  
summary(modelonormalizado2)  
summary(modelonormalizado3)  
summary(modelonormalizado4)
```

## 7. DECLARACIÓN DE USO DE HERRAMIENTAS DE INTELIGENCIA ARTIFICIAL GENERATIVA EN TRABAJOS DE FIN DE GRADO

**ADVERTENCIA:** Desde la Universidad consideramos que ChatGPT u otras herramientas similares son herramientas muy útiles en la vida académica, aunque su uso queda siempre bajo la responsabilidad del alumno, puesto que las respuestas que proporciona pueden no ser veraces. En este sentido, NO está permitido su uso en la elaboración del Trabajo fin de Grado para generar código porque estas herramientas no son fiables en esa tarea. Aunque el código funcione, no hay garantías de que metodológicamente sea correcto, y es altamente probable que no lo sea.

Por la presente, yo, Ana Bueno Rodríguez, estudiante de E2+ Business Analytics de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado "**Evaluación y Análisis del desempeño medioambiental en empresas altamente contaminantes y su impacto en el rendimiento financiero**", declaro que he utilizado la herramienta de Inteligencia Artificial Generativa ChatGPT u otras similares de IAG de código sólo en el contexto de las actividades descritas a continuación [el alumno debe mantener solo aquellas en las que se ha usado ChatGPT o similares y borrar el resto. Si no se ha usado ninguna, borrar todas y escribir "no he usado ninguna"]:

1. **Brainstorming de ideas de investigación:** Utilizado para idear y esbozar posibles áreas de investigación.
2. **Crítico:** Para encontrar contra-argumentos a una tesis específica que pretendo defender.
3. **Referencias:** Usado conjuntamente con otras herramientas, como Science, para identificar referencias preliminares que luego he contrastado y validado.
4. **Metodólogo:** Para descubrir métodos aplicables a problemas específicos de investigación.
5. **Interpretador de código:** Para realizar análisis de datos preliminares.
6. **Estudios multidisciplinares:** Para comprender perspectivas de otras comunidades sobre temas de naturaleza multidisciplinar.
7. **Corrector de estilo literario y de lenguaje:** Para mejorar la calidad lingüística y estilística del texto.
8. **Sintetizador y divulgador de libros complicados:** Para resumir y comprender literatura compleja.
9. **Generador de datos sintéticos de prueba:** Para la creación de conjuntos de datos ficticios.

10. **Generador de problemas de ejemplo:** Para ilustrar conceptos y técnicas.
11. **Revisor:** Para recibir sugerencias sobre cómo mejorar y perfeccionar el trabajo con diferentes niveles de exigencia.
12. **Traductor:** Para traducir textos de un lenguaje a otro.

Afirmo que toda la información y contenido presentados en este trabajo son producto de mi investigación y esfuerzo individual, excepto donde se ha indicado lo contrario y se han dado los créditos correspondientes (he incluido las referencias adecuadas en el TFG y he explicitado para que se ha usado ChatGPT u otras herramientas similares). Soy consciente de las implicaciones académicas y éticas de presentar un trabajo no original y acepto las consecuencias de cualquier violación a esta declaración.

Fecha: 20 de junio 2024

Firma: \_\_\_\_Ana Bueno Rodríguez\_\_\_\_\_

## 8. BIBLIOGRAFÍA.

Agencia Europea de Medio Ambiente. *Greenhouse gases - data viewer*.  
<https://www.eea.europa.eu/data-and-maps/data/data-viewers/greenhouse-gases-viewer>

Åhman, M. (2022). Unlocking the “Hard to Abate” Sectors. *World Resources Institute*. [https://portal.research.lu.se/files/77011421/Policy\\_Brief\\_WRI\\_expert\\_perspective.pdf](https://portal.research.lu.se/files/77011421/Policy_Brief_WRI_expert_perspective.pdf)

Alshehhi, A., Nobanee, H., & Khare, N. (2018). The impact of sustainability practices on corporate financial performance: Literature trends and future research potential. *Sustainability*, 10(2), 494.  
<https://www.mdpi.com/2071-1050/10/2/494>

Bhattacharyya, R. (2022). Green finance for energy transition, climate action and sustainable development: overview of concepts, applications, implementation and challenges. *Green Finance*, 4(1), 1-35. <https://www.aimspress.com/article/id/61b72862ba35de53d6ee0d3d>

Burnett, R.D. y Hansen, D.R. (2008). Eco-efficiency: Defining a role for environmental cost management. *Accounting, Organizations and Society*, 33(6), 551-581. <https://doi.org/10.1016/j.aos.2007.06.002>

Busch, T. y Hoffmann, V.H. (2011). How Hot Is Your Bottom Line? Linking Carbon and Financial Performance. *Business & Society*, 50(2), 233-265.

Castro de Doens, L. (2016). Financiamento verde para el desarrollo sostenible. *Economía y desarrollo*, 156(1), 155-167. [http://scielo.sld.cu/scielo.php?pid=S0252-85842016000100011&script=sci\\_arttext&tlng=pt](http://scielo.sld.cu/scielo.php?pid=S0252-85842016000100011&script=sci_arttext&tlng=pt)

Clark, G. L., Feiner, A., & Viehs, M. (2015). From the stockholder to the stakeholder: How sustainability can drive financial outperformance. *Available at SSRN 2508281*.  
<https://formafutura.com/assets/documents/2015-University-Oxford-From-the-Stockholder-to-the-Stakeholder.pdf>

de Francia, J. C., & Ayerbe, C. G. (2006). Repercusión económica de la certificación medioambiental ISO 14001. *Cuadernos de gestión*, 6(1), 45-62. <https://www.redalyc.org/pdf/2743/274320225003.pdf>

Deloitte. *Company ESG score impact on market value*. Deloitte Switzerland. <https://www2.deloitte.com/content/dam/Deloitte/ch/Documents/financial-advisory/deloitte-ch-company-ESG-score-impact-on-market-value.pdf>

Durán, M. P., Calderón, E. P., & Montero, P. M. (2018). Ecoeficiencia y sus efectos sobre el desempeño económico de las empresas del Dow Jones Sustainability World Index 2016. *Prisma Social: revista de investigación social*, (22), 271-295. <https://dialnet.unirioja.es/descarga/articulo/6562967.pdf>

Edelenbosch, O. Y., Hof, A. F., van den Berg, M., de Boer, H. S., Chen, H. H., Daioglou, V., ... & van Vuuren, D. P. (2024). Reducing sectoral hard-to-abate emissions to limit reliance on carbon dioxide removal. *Nature Climate Change*, 1-8. <https://www.nature.com/articles/s41558-024-02025-y>

Fu, T., & Li, J. (2023). An empirical analysis of the impact of ESG on financial performance: the moderating role of digital transformation. *Frontiers in Environmental Science*, 11, 1256052. <https://www.frontiersin.org/articles/10.3389/fenvs.2023.1256052/pdf>

Galeano, V. J. C., & Rubia, B. N. (2022). La captura de CO2. Un pilar indispensable de la descarbonización. In *Energía y Geoestrategia 2022* (pp. 197-294). Instituto Español de Estudios Estratégicos. <https://dialnet.unirioja.es/descarga/articulo/8536312.pdf>

Grábalos, E. F., Lafuente, P. P., y Bahillo, C. G. (2017). Desarrollo humano desde la perspectiva del crecimiento. *Revista de Ciencias Sociales (Ve)*, 23(4), 81-97.

Gompers, P., Ishii, J., & Metrick, A. (2003). Corporate governance and equity prices. *The quarterly journal of economics*, 118(1), 107-156.

Hart, S.L. (1995). A Natural-Resource-Based view of the firm. *Academy of Management Review*, 20(4), 986-1014. doi: <https://doi.org/10.5465/amr.1995.9512280033>

Hart, S.L. y Ahuja, G. (1996). Does it pay to be green? An empirical examination of the relationship between emission reduction and firm performance. *Business Strategy and the Environment*, 5(1), 30-37. [https://doi.org/10.1002/\(SICI\)1099-0836\(199603\)5:1<30::AID-BSE38>3.0.CO;2-Q](https://doi.org/10.1002/(SICI)1099-0836(199603)5:1<30::AID-BSE38>3.0.CO;2-Q)

Herrera, E. A., de la Garza Cárdenas, M. H., & Mendoza, F. J. D. (2024). Prácticas sostenibles en el desempeño financiero de empresas mexicanas. *Revista del Centro de Investigación de la Universidad la Salle*, 16(61), 201-224. <https://revistasinvestigacion.lasalle.mx/index.php/recein/article/download/3866/3640>

Klassen, R.D. y McLaughlin, C.P. (1996). The impact of environmental management on firm performance. *Management Science*, 42(8), 1199-1214. <https://doi.org/10.1287/mnsc.42.8.1199>

Madueño, J. H., Jorge, M. L., & Martínez-Martínez, D. (2013). Relación entre responsabilidad social y performance en las pequeñas y medianas empresas: Revisión bibliográfica. *Cuadernos de gestión*, 13(2), 39-65. <https://www.redalyc.org/pdf/2743/274326464006.pdf>

McKinsey & Company. (2022). *Hacia un futuro sostenible, inclusivo y en crecimiento: El papel de las empresas*. <https://www.mckinsey.com/featured-insights/destacados/hacia-un-futuro-sostenible-inclusivo-y-en-crecimiento-el-papel-de-las-empresas/es>

Moors, E.H.M., Mulder, K.F. y Vergragt, P.J. (2005). Towards cleaner production: barriers and strategies in the base metals producing industry. *Journal of Cleaner Production*, 13(7), 657-668. <https://doi.org/10.1016/j.jclepro.2003.12.010>

Murillo, J.L. (2005). Determinantes del comportamiento estratégico medioambiental de las empresas industriales de Aragón (Tesis Doctoral). Universidad de Zaragoza, España.

Murillo-Luna, J. L., Garcés-Ayerbe, C. y Rivera-Torres, P. (2007). What prevents firms from advancing in their environmental strategy? *International Advances in Economic Research*, 13(1), 35-46. <https://doi.org/10.1007/s11294-006-9059-6>

NASA. *Global temperature: Vital signs of the planet*. NASA. <https://climate.nasa.gov/vital-signs/global-temperature/?intent=121>

Paltsev, S., Morris, J., Kheshgi, H., & Herzog, H. (2021). Hard-to-Abate Sectors: The role of industrial carbon capture and storage (CCS) in emission mitigation. *Applied Energy*, 300, 117322. <https://www.sciencedirect.com/science/article/pii/S0306261921007327>

Pérez, P. (2024, 4 marzo). ¿Cuántas empresas certificadas en ISO 14001 hay en el mundo? - Nueva ISO 14001. Nueva ISO 14001. <https://www.nueva-iso-14001.com/2024/03/cuantas-empresas-certificadas-en-iso-14001-hay-en-el-mundo/>

Porter, M.E. (1991). America's green strategy. *Scientific American*, 264(4), 168-179.

Porter, M. E., & Linde, C. V. D. (1995). Toward a new conception of the environment-competitiveness relationship. *Journal of economic perspectives*, 9(4), 97-118. <https://pubs.aeaweb.org/doi/pdf/10.1257/jep.9.4.97>

Roffé, M. A., & Ignacio González, F. A. (2024). El impacto de las prácticas sostenibles en el desempeño financiero de las empresas: Una revisión de la literatura. *Visión de futuro*, 28(1), 195-220. [http://www.scielo.org.ar/scielo.php?pid=S1668-87082024000100195&script=sci\\_arttext](http://www.scielo.org.ar/scielo.php?pid=S1668-87082024000100195&script=sci_arttext)

Russo, M.V. y Fouts, P.A. (1997). A Resource Based Perspective on Corporate Environmental Performance and Profitability. *Academy of Management Journal*, 40(3), 534-559. <https://doi.org/10.5465/257052>

Sachs, J. D., Woo, W. T., Yoshino, N., & Taghizadeh-Hesary, F. (2019). Importance of green finance for achieving sustainable development goals and energy security. *Handbook of green finance*, 3.

Sharma S. y Vredenburg H. (1998). Proactive corporate environmental strategy and the development of competitively valuable organizational capabilities. *Strategic Management Journal*, 19(8), 729-753.

Shrivastava, P. (1995). The Role of Corporations in Achieving Ecological Sustainability. *Academy of Management Review*, 20(4), 936-960. <https://doi.org/10.5465/amr.1995.9512280026>

Valenzuela Fernández, L., Jara-Bertin, M., & Villegas Pineaur, F. (2015). Prácticas de responsabilidad social, reputación corporativa y desempeño financiero. *Revista de Administração de Empresas*, 55, 329-344. <https://www.scielo.br/j/rae/a/SpfJybqrX7DhGk4mXh9twXh/?lang=es>

Zhang, C., Zhou, B., Wang, Q., & Jian, Y. (2024). The consequences of environmental big data information disclosure on hard-to-abate Chinese enterprises' green innovation. *Journal of Innovation & Knowledge*, 9(2), 100474. <https://www.sciencedirect.com/science/article/pii/S2444569X24000143>