



FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES  
GRADO EN ADMINISTRACIÓN Y DIRECCIÓN DE EMPRESAS

Trabajo Fin de Grado

*¿Es imparable el fenómeno de las Finanzas Sostenibles?*

Análisis de su evolución y del posible impacto en la cotización de Repsol a través de técnicas de *Data Analytics*.

Autor:

Carlos Varona Vidaurrezaga

Directora:

María Coronado Vaca

Madrid, junio 2021

## **Resumen**

El calentamiento global es un hecho que hoy en día podríamos decir que es empírico y que, aunque discutido, no es puesto en duda por ningún sector socioeconómico. La concienciación en todos los ámbitos es creciente y el mundo empresarial y financiero es actor cada vez más relevante en este proceso de adaptación de uso de la energía que está siendo promovido, impulsado y supervisado por los distintos organismos internacionales.

Dicho proceso creciente ha sido objeto de análisis en este estudio mediante técnicas de *Data Analytics*, utilizando, por un lado, la base de datos de *New York Times* para ver y analizar la evolución de las menciones en artículos del *New York Times* de distinta terminología sostenible.

Por otro lado, el trabajo ha abordado el estudio sobre el impacto que la creciente concienciación sobre la sostenibilidad ha podido tener en la cotización de una empresa sensible a sus riesgos. Por ello, se ha escogido Repsol, que está considerada como una empresa altamente expuesta a cuestiones medioambientales, de transición energética y de movilidad, y se ha calculado el impacto de la sostenibilidad en su cotización bursátil a través de los sentimientos extraídos en *tweets* sobre Repsol y la sostenibilidad.

Para la realización de los mencionados análisis se ha hecho uso de diferentes técnicas mediante los softwares de programación *Rstudio* y *Python*.

El análisis de la base de datos del NYT sobre la evolución de las menciones de terminología sostenible ha mostrado unos resultados dispares y que no refutan al completo el auge experimentado por las finanzas sostenibles.

Con relación al estudio sobre el posible impacto de los sentimientos en los movimientos bursátiles de la empresa petrolífera, se ha concluido que no existe correlación entre el precio de cotización de Repsol y los sentimientos de Twitter respecto a su gestión sostenible.

## **Palabras clave**

Finanzas sostenibles, sostenibilidad, Repsol, cotización, Twitter, análisis de sentimientos, webscraping.

## **Abstract**

Global warming is a fact that today we could say is empirical and that, although disputed, is not questioned by any socioeconomic sector. Awareness in all areas is growing and the business and financial world is an increasingly relevant actor in this process of adaptation of energy use that is being promoted, driven and supervised by various international organizations.

This growing process has been analyzed in this study by means of *Data Analytics* techniques, using, on the one hand, the *New York Times* database to see and analyze the evolution of mentions in *New York Times* articles of different sustainable terminology.

On the other hand, the work has addressed the study on the impact that the growing awareness of sustainability has had on the share price of a company sensitive to its risks. For this reason, Repsol, which is considered a company highly exposed to environmental, energy transition and mobility issues, was chosen and the impact of sustainability on its stock price was calculated through the sentiments extracted from *tweets* about Repsol and sustainability.

In order to carry out the aforementioned analyses, different techniques were used through the programming software *Rstudio* and *Python*.

The analysis of the NYT database on the evolution of mentions of sustainable terminology has shown disparate results that do not fully refute the reality of the boom experienced by sustainable finance.

With regard to the study on the possible impact of sentiment on the oil company's stock market movements, it was concluded that there is no correlation between Repsol's share price and Twitter's sentiment regarding its sustainable management.

## **Keywords**

Green Finance, sustainability, Repsol, quote, Twitter, sentimental analysis, webscraping.

## ÍNDICE DE ILUSTRACIONES

<i>Ilustración 1: Evolución de la temperatura media anual en tierra y océanos desde el año 1880 hasta 2015.....</i>	<i>11</i>
<i>Ilustración 2: Incremento de la temperatura global respecto a la era industrial.....</i>	<i>11</i>
<i>Ilustración 3: Desastres anuales con desplazamiento de personas. ....</i>	<i>12</i>
<i>Ilustración 4: Captura de los ODS más apoyados por Repsol.....</i>	<i>14</i>
<i>Ilustración 5: Los 17 Objetivos de Desarrollo Sostenible (ODS).....</i>	<i>20</i>
<i>Ilustración 6: Emisiones anuales de deuda sostenible a nivel mundial, en miles de millones de \$ en el periodo 2012 a 2018.....</i>	<i>23</i>
<i>Ilustración 7: Extracto del código usado en RStudio para la descarga de las cotizaciones de Repsol en el MCE .....</i>	<i>40</i>
<i>Ilustración 8: Extracto del código usado en Python para la descarga de tweets .....</i>	<i>41</i>
<i>Ilustración 9: Llamada a la Api key .....</i>	<i>43</i>
<i>Ilustración 10: Funciones para la base URL y el querying .....</i>	<i>43</i>
<i>Ilustración 11: Funciones para establecer intervalos temporales .....</i>	<i>44</i>
<i>Ilustración 12: Funciones para contar las apariciones de un término en un intervalo de tiempo determinado. ....</i>	<i>44</i>
<i>Ilustración 13: Funciones para contar las apariciones de un término en un intervalo de tiempo determinado y para crear un gráfico con su evolución.....</i>	<i>45</i>
<i>Ilustración 14: Evolución de las menciones de Green Finance por año en el NYT.....</i>	<i>47</i>
<i>Ilustración 15: Evolución de las menciones de Green Finance por mes en el NYT. ....</i>	<i>48</i>
<i>Ilustración 16: Evolución de las menciones de Sustainable Finance por año en el NYT. ....</i>	<i>50</i>
<i>Ilustración 17: Evolución de las menciones de Sustainable Finance por mes en el NYT. ....</i>	<i>50</i>
<i>Ilustración 18: Evolución de las menciones de Sustainability por año en el NYT.....</i>	<i>51</i>
<i>Ilustración 19: La evolución de las menciones de Sustainability por mes en el NYT....</i>	<i>51</i>
<i>Ilustración 20: La evolución de las menciones de Green Bonds por año en el NYT. ....</i>	<i>52</i>
<i>Ilustración 21: Evolución de las menciones de Green Bonds por mes en el NYT. ....</i>	<i>53</i>
<i>Ilustración 22: Evolución de las menciones de Green Deal por año en el NYT.....</i>	<i>54</i>
<i>Ilustración 23: Evolución de las menciones de Green Deal por mes en el NYT.....</i>	<i>55</i>
<i>Ilustración 24: Nube de palabras con mayor presencia tras haber eliminado las stopwords en los tweets analizados de Repsol. ....</i>	<i>56</i>

<i>Ilustración 25: Regla para calcular el número óptimo de temas o clusters .....</i>	<i>57</i>
<i>Ilustración 26: Temas y sus palabras más descriptivas .....</i>	<i>57</i>
<i>Ilustración 27: Frecuencia de sentimientos .....</i>	<i>58</i>
<i>Ilustración 28: Gráfico de barras para ver las emociones con mayor presencia.....</i>	<i>59</i>
<i>Ilustración 29: Nubes de emociones.....</i>	<i>60</i>
<i>Ilustración 30: Evolución de sentimientos en los tweets, siendo Syuzhet Plot la gráfica con todas las medidas del algoritmo y Simplified Macro Shape el gráfico normalizado del anterior. ....</i>	<i>61</i>
<i>Ilustración 31: Precio de cierre ajustado de las acciones de Repsol entre 2015 y 2020. ....</i>	<i>62</i>
<i>Ilustración 32: Extracto de los ratings de Sustainalytics sobre Repsol en relación con su calificación sobre los riesgos ESG.....</i>	<i>64</i>
<i>Ilustración 33: Nivel de controversia de Repsol según los criterios de Sustainalytics..</i>	<i>64</i>
<i>Ilustración 34: Extracto de los ratings de Sustainalytics sobre Repsol en materia de exposición al riesgo y gestión del riesgo.....</i>	<i>65</i>
<i>Ilustración 35: Correlación entre el sentimiento promedio semanal de los tweets y el precio de cierre ajustado semanal en el intervalo entre 2015 y 2020. ....</i>	<i>66</i>

## LISTA DE ACRÓNIMOS

API	<i>Application programming interface</i>
CBI	<i>Climate Bonds Initiative</i>
CE	Comisión Europea
COP21	Conferencia de las Naciones Unidas sobre Cambio Climático 2015
ESG	<i>Environmental, Social and Governance</i>
FSB	Consejo de Estabilidad Financiera
FTJ	Fondo de Transición Justa
GBP	<i>Green Bonds Principles</i>
GEI	Gases Efecto Invernadero
ICMA	Asociación Internacional de Mercados de Capitales
ISR	Inversiones Socialmente Responsables
MDPI	<i>Multidisciplinary Digital Publishing Institute</i>
MSCI	Índice bursátil <i>Morgan Stanley Capital International</i>
NDC	Contribuciones Nacionalmente Determinadas
NFRD	<i>Non-Financial Reporting Directive</i>
NGFS	<i>Network for Greening the Financial System</i>
ODM	Objetivos de Desarrollo del Milenio
ODS	Objetivos de Desarrollo Sostenible
OECD	Organización para la Cooperación y Desarrollo Económico
OMS	Organización Mundial de la Salud

PACTA	<i>Paris Agreement Capital Transition Assessment</i>
PAI	<i>Principal Adverse Impact</i>
PCAF	Partnership for Carbon Accounting Financials
SFDR	<i>Sustainable Finance Disclosure Regulation</i>
TCFD	Task Force on Climate-related Financial Disclosures
TEG	Grupo Técnico de Expertos
UE	Unión Europea
URL	<i>Uniform Resource Locator</i>

## ÍNDICE DE CONTENIDOS

<b>1. INTRODUCCIÓN</b> .....	<b>10</b>
<b>1.1. Objetivos generales y específicos</b> .....	<b>10</b>
<b>1.2. Justificación del tema objeto de estudio</b> .....	<b>10</b>
<b>1.3. Metodología y estructura.</b> .....	<b>15</b>
<b>2. NOCIONES BÁSICAS SOBRE LAS FINANZAS SOSTENIBLES</b> .....	<b>16</b>
<b>2.1. Finanzas sostenibles: definición y ámbitos</b> .....	<b>16</b>
<b>2.2. Evolución de la Finanzas Sostenibles</b> .....	<b>17</b>
2.2.1. Acuerdo de París.....	18
2.2.2. Cumbre de Desarrollo Sostenible: Agenda 2030 y ODS .....	19
<b>2.3. Otras iniciativas de desarrollo de finanzas sostenibles</b> .....	<b>20</b>
2.3.1. Iniciativas desde el sector público .....	21
2.3.2. Iniciativas desde el sector privado.....	29
<b>2.4. Criterios ASG</b> .....	<b>33</b>
<b>2.5. La pandemia y la sostenibilidad</b> .....	<b>33</b>
2.5.1. NextGenerationEU .....	35
<b>3. CONCEPTOS BÁSICOS DE TEXT MINING, WEBSCRAPING Y ANÁLISIS DE SENTIMIENTOS</b> .....	<b>36</b>
<b>4. METODOLOGÍA</b> .....	<b>38</b>
<b>4.1. Descarga de datos</b> .....	<b>38</b>
4.1.1. Descarga de datos del New York Times .....	38
4.1.2. Descarga de datos de las cotizaciones de Repsol. ....	39
4.1.3. Descarga de datos de Twitter.....	40
<b>4.2. Análisis exploratorio y de sentimiento</b> .....	<b>42</b>
4.2.1. Menciones en artículos del New York Times .....	42
4.2.2. Análisis de sentimiento.....	45
<b>5. RESULTADOS</b> .....	<b>47</b>
<b>5.1. Resultados de la evolución de la terminología en el NYT.</b> .....	<b>47</b>
<b>5.2. Resultados del análisis de sentimiento de los tweets relativos a Repsol ..</b>	<b>56</b>
<b>5.3. Análisis del precio de cotización de Repsol</b> .....	<b>62</b>



5.4. Análisis de correlación entre sentimiento promedio de los tweets y el precio de cotización de Repsol.....	65
6. CONCLUSIONES .....	68
6.1. Limitaciones del estudio .....	70
7. BIBLIOGRAFÍA .....	72
8. ANEXOS.....	77
8.1. Código utilizado para el análisis de la evolución de la terminología sostenible en cuanto a menciones en artículos del NYT.....	77
8.2. Código utilizado para la descarga de las cotizaciones de Repsol y para la elaboración gráfico que muestra su evolución.....	81
8.3. Código utilizado en Python para la descarga de los tweets sobre Repsol y Sostenibilidad en el período que va de enero de 2015 hasta diciembre de 2020.	81
8.4. Códigos utilizados en RStudio para la realización del análisis de sentimiento de los tweets descargados previamente.....	82
8.5. Código utilizado de RStudio para obtener le gráfico con la evolución del sentimiento de los tweets .....	95

*“Si un intercambio entre dos partes es voluntario, no tendrá lugar a menos que ambas partes crean que se beneficiarán de él”*

*Milton Friedman*

## **1. INTRODUCCIÓN**

### **1.1. Objetivos generales y específicos**

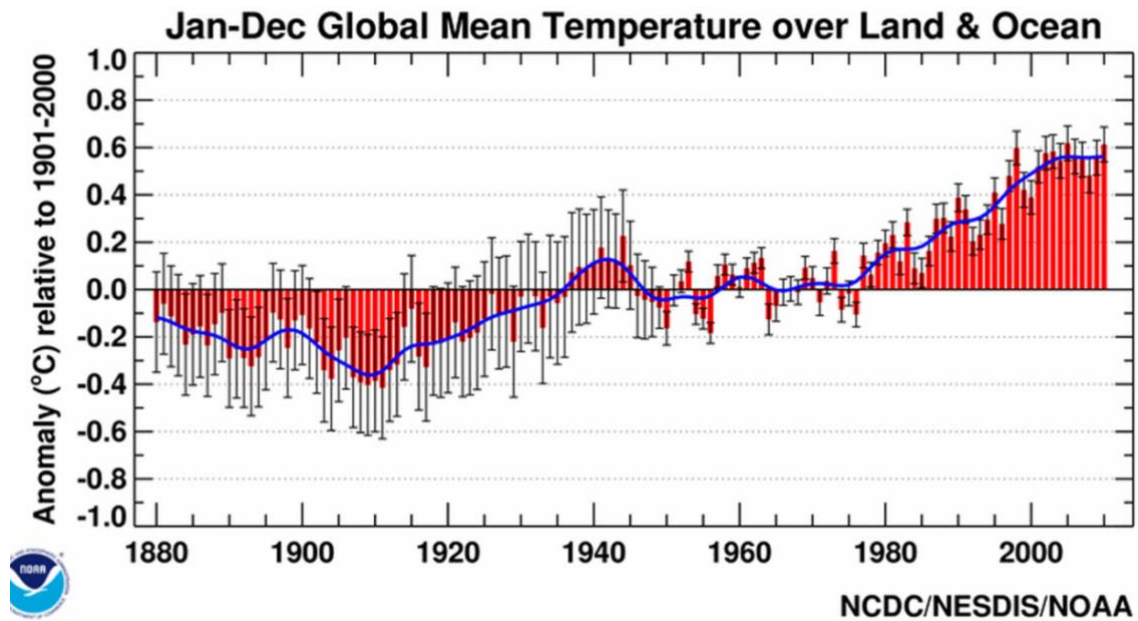
El presente trabajo de investigación contiene tres principales objetivos. El primero es presentar la evolución que han tenido las finanzas sostenibles y su perspectiva de crecimiento en los próximos años, mencionando las diferentes iniciativas tanto desde el sector público como del privado, y analizando la evolución de la terminología sostenible en un medio de comunicación prestigioso como el New York Times. El segundo es ver sobre la cotización de Repsol el posible impacto que ha tenido la sostenibilidad en sus movimientos, además de, a través de diferentes metodologías de *data analytics*, analizar el sentimiento de Twitter sobre la conducta de Repsol en términos de sostenibilidad. Por último, se estudiará si existe correlación entre la cotización de Repsol y los sentimientos extraídos de Twitter.

### **1.2. Justificación del tema objeto de estudio**

Hasta hace unas escasas décadas, los movimientos ecologistas ocupaban unos espacios políticos y mediáticos de muy escasa relevancia, quedando ceñida su presencia a algunos países de alto nivel de desarrollo económico como, por ejemplo, Alemania.

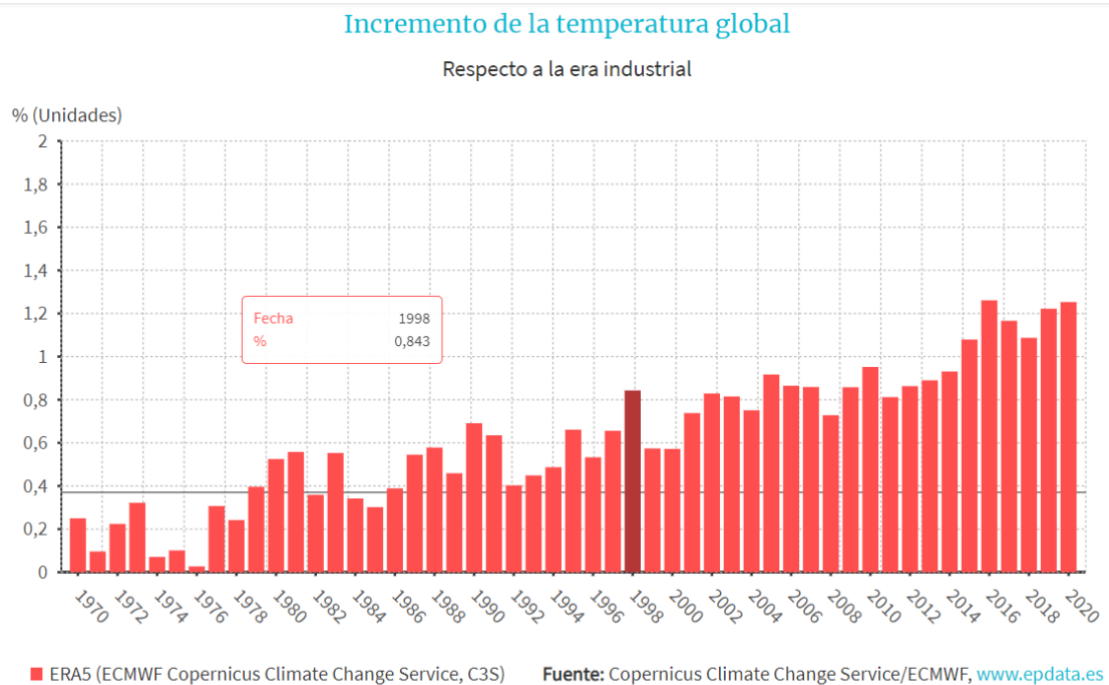
De forma concurrente con lo anterior, la comunidad científica en su afán investigador sobre todas las cuestiones relativas al planeta y su naturaleza comprobó con cierta preocupación como se producía de forma exponencial el calentamiento global. Este hecho científico se ha podido demostrar con la medición de varias magnitudes entre las que se encuentran la temperatura media del planeta (visible en las Ilustraciones 1 y 2), la dimensión de la capa de ozono o el derretimiento de los casquetes polares.

**Ilustración 1: Evolución de la temperatura media anual en tierra y océanos desde el año 1880 hasta 2015.**



Fuente: National Oceanic and Atmospheric Administration (NOAA)

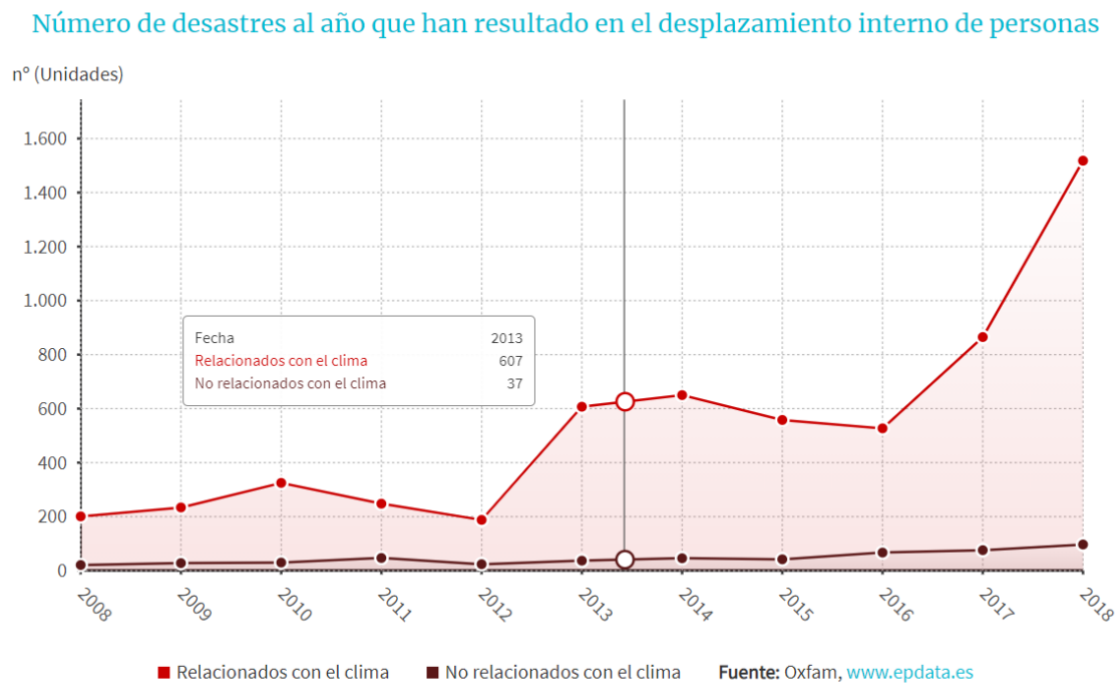
**Ilustración 2: Incremento de la temperatura global respecto a la era industrial.**



Fuente: Copernicus Climate Change Service Copernicus Climate Change Service

Otro hecho irrefutable que ha removido también la conciencia de todos los legisladores y de la población en general ha sido la frecuente aparición de fenómenos meteorológicos con efectos devastadores (evolución visible en Ilustración 3), como riadas en cualquier zona del planeta con independencia del nivel de dotaciones de infraestructuras del país afectado.

**Ilustración 3: Desastres anuales con desplazamiento de personas.**



Fuente: Oxfam, [www.epdata.es](http://www.epdata.es)

A nadie se le escapa que la incorporación de países fuertemente poblados como China y la India al desarrollo económico, al que todo pueblo tiene perfecto derecho, ha traído en parte como consecuencia la aceleración de los procesos de calentamiento global.

Por si lo anterior no fuera suficiente, asistimos como usuarios de energía eléctrica a todos los niveles (industrial y doméstico) a la influencia cada vez más notable en el precio de la energía de los certificados de emisión de CO<sub>2</sub> y que sin duda harán más atractivos los activos financieros verdes para un público inversor cada vez más amplio.

Una vez descrita brevemente la situación que existe en el planeta actualmente como consecuencia de un comportamiento ajeno a las cuestiones medioambientales con efectos cada vez más patentes y visibles, que ha provocado la aparición de normativa y acuerdos

internacionales y de instrumentos financieros para vehicular mejor las inversiones en proyectos de eficiencia energética e infraestructuras sostenibles, se ha querido estudiar el impacto en la opinión pública del “Green Finance” mediante la utilización de las herramientas informáticas que el *Big Data* ha puesto a nuestra disposición. El análisis mediante las técnicas de *Business Analytics* se ha centrado en los siguientes temas:

- La evolución del impacto de las finanzas sostenibles en un periódico prestigioso de amplia tirada como el NYT<sup>1</sup>.
- El análisis del potencial impacto que la aparición de la Agenda 2030 y otros acuerdos y hechos relevantes sobre la sostenibilidad han tenido en la cotización de Repsol mediante un análisis de sentimientos en Twitter.
- Un análisis de correlación entre los movimientos bursátiles de la empresa petrolífera y el sentimiento promedio extraído de los *tweets* sobre la propia empresa y la sostenibilidad.

La elección de la base de datos del NYT se debe a la gran repercusión que tiene este histórico medio de comunicación en todo el mundo, especialmente desde la década de 1970, que coincide a su vez con los primeros acontecimientos relevantes en materia sostenible. Fue creado en 1851 y es considerado por muchos el diario de referencia en los Estados Unidos, siendo galardonado con premios de alto prestigio como el *Pullitzer*, que ha ganado en más de 130 ocasiones en los últimos cien años y que premia logros en el periodismo impreso y *online*. Además, su base de datos, con acceso gratuito, recoge información con un intervalo temporal muy amplio y con unos límites que no afectan a nuestro estudio.

Para realizar el análisis de sentimientos y de correlación comentado anteriormente, se decidió utilizar la información presentada en Twitter. Se trata de una red social que cuenta con más de 340 millones de usuarios únicos mensuales, de los que más del 70% la usa como fuente de información, en la que se puede compartir cualquier tipo de información u opinión de forma rápida, gratuita y simple (Onmicore, 2021). Esto posibilita que, a través de técnicas de *Business Analytics*, se puedan recoger sentimientos de los *tweets* enviados en su plataforma, permitiendo analizar el posible impacto de la gestión sostenible percibida por el público general de Repsol en su cotización bursátil.

---

<sup>1</sup>NYT son las siglas del medio de comunicación New York Times

Además, otro motivo que ha suscitado la voluntad de utilizar los datos de Twitter es que se considera que está dejando de lado a los medios de comunicación más tradicionales, por lo que puede ser una fuente interesante para ampliar la información que puede brindarnos un medio de comunicación tradicional como el NYT.

Se ha considerado procedente elegir una empresa como Repsol por su marcado carácter industrial en un sector fuertemente impactado por la transición energética en la que todos estamos inmersos y vamos a experimentar en los próximos años, y muy preocupada por su imagen de compañía seria y responsable en todas las cuestiones medioambientales. Es destacable el hecho de que Repsol haya sido la primera compañía energética en asumir el compromiso de alcanzar cero emisiones netas en 2050 en línea con los objetivos climáticos del Acuerdo de París y los Objetivos de Desarrollo Sostenible de Naciones Unidas. El más claro exponente de lo anterior es la existencia (disponible en la página web oficial de Repsol) de un Plan Global de Sostenibilidad, de una Hoja de Ruta ante el Cambio Climático y un Informe ODS<sup>2</sup> 2020 propios de la empresa.

***Ilustración 4: Captura de los ODS más apoyados por Repsol***



*Fuente: Hoja de Ruta de Cambio Climático de Repsol (Top-Priority SDGs), extraído de la página web oficial de Repsol.*

Para utilizar ambas bases de datos (NYT y Twitter) se va a hacer uso de diferentes técnicas de *data analytics*, que puedan ofrecernos información relevante sobre el fenómeno de las finanzas sostenibles. En primer lugar, el *webscraping* para transformar la información de las bases de datos en datos estructurados, de forma que puedan ser tratados para su análisis. El *Text Mining*, por su parte, nos permitirá trabajar con los datos y extraer información de ellos. Por último, la realización de un análisis de sentimientos

<sup>2</sup> ODS son las siglas de los Objetivos de Desarrollo Sostenible

de los *tweets* sobre Repsol y la sostenibilidad será de utilidad para extraer los sentimientos y emociones que muestran los *tweets* y así poder ver lo que suscita la gestión sostenible de la empresa española, además de poder analizar la posible correlación entre el sentimiento promedio de los *tweets* y la cotización bursátil de Repsol.

### **1.3. Metodología y estructura.**

La metodología del estudio ha sido la siguiente. En primer lugar, se ha realizado una metodología general de búsqueda de información utilizando diferentes fuentes de información de gran peso como el Banco de España, la Unión Europea o diferentes informes de foros especializados en el tema que nos concierne. En segundo lugar, una metodología empírica, que se desarrolla mediante el uso de las herramientas que nos brinda el *data analytics*.

La estructura del trabajo de investigación se divide en cinco partes, obviando el presente capítulo 1. El capítulo 2 se centra en presentar la metodología general, de forma que sirva de contexto para entender el concepto de las finanzas sostenibles y los diferentes acuerdos que han apoyado su evolución y crecimiento. A continuación, en el capítulo 3 se exponen diferentes nociones básicas sobre el *text mining*, el *webscraping* y el *sentimental analysis*, herramientas fundamentales para la elaboración de la metodología empírica, que será plasmada en los capítulos 4 y 5, siendo el primero de ellos el encargado de informar sobre la metodología utilizada para realizar la descarga de datos y el posterior análisis exploratorio. El capítulo 5, por su parte, presenta los resultados e interpretaciones obtenidas, incluyendo la posible correlación del sentimiento promedio de Twitter con los movimientos en el precio ajustado de cotización de Repsol. Finalmente, el capítulo 6 se centra en las conclusiones del trabajo y en las limitaciones que se han ido presentando a lo largo de su elaboración, al objeto de que puedan ser consideradas en otros estudios en el campo de las finanzas sostenibles y el *data analytics*.

## 2. NOCIONES BÁSICAS SOBRE LAS FINANZAS SOSTENIBLES

### 2.1. Finanzas sostenibles: definición y ámbitos

El cambio climático es uno de los mayores retos que tiene que afrontar la sociedad actual por sí misma para que no se ponga en riesgo ni se comprometa la capacidad de las generaciones futuras. Se trata de un problema universal que afecta a todos, pero cuya intensidad varía según el país, habiendo unos con mayor exposición al cambio. Por ejemplo, España entra dentro de los 40 países más expuestos a hechos climáticos según el Índice de Riesgo Climático<sup>3</sup>.

Ante esta situación en la que la sostenibilidad se trata como un tema central para multitud de países, el sistema financiero tiene un papel fundamental. Que sus agentes tomen conciencia y aumenten las inversiones que sean consecuentes con los objetivos sostenibles comunes es primordial para que estos objetivos se cumplan. Estas inversiones pueden ser denominadas ISR<sup>4</sup> (Inversiones Socialmente Responsables), que son un tipo de inversión que se aleja de las centradas únicamente en la rentabilidad económica, buscando ahora también un impacto social o medioambiental.

El término Finanzas Sostenibles no ha nacido ahora, pero es un fenómeno imparable que está tomando gran relevancia en el último lustro gracias a diferentes acuerdos que serán presentados a lo largo del capítulo presente. Además, los “*millennials*” están liderando diferentes preocupaciones referentes a la sostenibilidad, lo que propicia aún en mayor medida el auge del término (World Economic Forum, 2017).

Las finanzas sostenibles carecen de una definición universal comúnmente aceptada, pero las podemos describir como:

"la financiación para realizar inversiones teniendo en cuenta cuestiones medioambientales, sociales y de gobierno" (Comisión Europea, 2020c).

---

<sup>3</sup> España ocupa el puesto 32 entre los países más amenazados por los eventos extremos (Germanwatch, 2021).

<sup>4</sup> ISR son las siglas de Inversión Socialmente Responsable.



Por tanto, podemos ver que las finanzas sostenibles pretenden potenciar el crecimiento económico sin descuidar variables como el medioambiente, actuando con una mayor transparencia y sensibilidad para luchar contra los riesgos relacionados con la sostenibilidad contando con el apoyo de los distintos actores financieros.

Las finanzas sostenibles abarcan todos los ámbitos del sistema financiero: bancos, gestores de activos y fondos, aseguradoras y demás organizaciones e individuos que tienen actividad financiera. Además, no solo hay que ver la sostenibilidad en su aspecto medioambiental, sino también en términos sociales (diversidad, inclusión, de no discriminación por motivos de género, etnia, religión, orientación sexual, etc.) y en términos de gobierno (diversidad en los consejos de administración, promoción de la salud y el bienestar de los trabajadores, condiciones laborales justas, etc.).

## **2.2. Evolución de la Finanzas Sostenibles**

Como se ha presentado en el punto anterior, el término Finanzas Sostenibles no es nuevo, ya que las primeras iniciativas datan de la década de los 70, pero ha cobrado mayor importancia en los últimos años, especialmente tras acuerdos como los Objetivos de Desarrollo del Milenio, el Protocolo de Kioto, los Acuerdos de Paris o la Agenda 2030 y los ODS que serán analizados, entre otros, en el presente capítulo 2.

Los ODM<sup>5</sup> de las Naciones Unidas son ocho objetivos acordados por los 192 Estados Miembros de las ONU<sup>6</sup> que nacieron con la intención de ser alcanzados en el 2015. La Declaración del Milenio de las Naciones Unidas firmada en septiembre de 2000, compromete a los dirigentes mundiales a luchar contra la pobreza, el hambre, la enfermedad, el analfabetismo, la degradación del medio ambiente y la discriminación contra la mujer. Los ODM que dimanaron de esa Declaración, tienen metas e indicadores específicos (Organización Mundial del Comercio, s. f.).

Los ocho Objetivos de Desarrollo del Milenio son:

---

<sup>5</sup> ODM se refiere a los Objetivos de Desarrollo del Milenio

<sup>6</sup> ONU son las siglas de la Organización Naciones Unidas

- 1- Erradicar la pobreza extrema y el hambre.
- 2- lograr la enseñanza primaria universal.
- 3- promover la igualdad entre los sexos y la autonomía de la mujer.
- 4- reducir la mortalidad infantil.
- 5- mejorar la salud materna.
- 6- combatir el VIH/SIDA, el paludismo y otras enfermedades.
- 7- garantizar la sostenibilidad del medio ambiente.
- 8- fomentar una asociación mundial para el desarrollo.

Este acuerdo provocó un punto de inflexión, estableciendo unos objetivos concretos y que contaban con el respaldo de los 192 países involucrados, lo que se tradujo en un intento de equilibrio a nivel global.

El Protocolo de Kioto de 1997 supuso también un hito importante en la concienciación de los dirigentes políticos sobre la sostenibilidad del Medio Ambiente. Fue propuesto en el seno de la Convención Marco de las Naciones Unidas sobre el Cambio Climático y alentaba a reducir en un 5% la emisión de gases de efecto invernadero (GEI<sup>7</sup>) por parte de 37 de los países más industrializados, con relación al año 1990. La Convención exhortaba al cumplimiento de metas vinculantes mediante la adopción de políticas y medidas de mitigación, con informes periódicos, aunque se tuvo que prorrogar la fecha fijada para su cumplimiento del 2012 al 2020 (Manos Unidas, 2020).

### **2.2.1. Acuerdo de París**

El Acuerdo de París es un tratado internacional que se aprobó el 15 de diciembre de 2015, entrando en vigor el 4 de noviembre de 2016, y que fue acordado por 195 países en la COP21<sup>8</sup>. Este acuerdo sustituyó al Protocolo de Kioto. Se trata del primer acuerdo vinculante jurídicamente entre los 195 países firmantes, que lo convierten en un hito en la historia de la normativa sostenible.

---

<sup>7</sup> GEI son las siglas por las que se conoce a los Gases de Efecto Invernadero

<sup>8</sup> La COP21 se refiere a la Conferencia de las Naciones Unidas sobre el Cambio Climático de 2015.

Principalmente, este acuerdo (Acuerdo de París) se basa en un plan de actuación para combatir el cambio climático y limitar el calentamiento del clima por debajo de 2 grados centígrados, además de luchar para reducir la emisión de gases de efecto invernadero. El Acuerdo incluye una serie de objetivos en términos de mitigación, adaptación y resiliencia que son aceptados por los países firmantes y en los que los mercados financieros son clave para su cumplimiento, estableciéndose un ciclo de revisión cada 5 años a partir de 2023 sobre el estado de implementación del objetivo del límite de los 2 grados. En 2020 se presentaron por parte de los países los planes de acción, conocidos como contribuciones determinadas a nivel nacional (NDC<sup>9</sup>). En estos planes de acción, cada país presenta las medidas que adoptará para el cumplimiento de los objetivos del Acuerdo de París, y establecen, a su vez, la manera en la que van a adaptarse a las consecuencias del calentamiento global.

### **2.2.2. Cumbre de Desarrollo Sostenible: Agenda 2030 y ODS**

En el mismo año en el que se produjo la aprobación del Acuerdo de París, 2015, tuvo lugar en Naciones Unidas la Cumbre de Desarrollo Sostenible. En esta cumbre, en la que estuvieron presentes gran cantidad de gobiernos y jefes de Estado, se aprobó tanto la Agenda 2030 como los Objetivos de Desarrollo Sostenible.

Los ODS son 17 objetivos (visibles en la Ilustración 5) y 169 metas que fueron aceptados en la Cumbre mencionada, y que, además, contienen una declaración en la que se presentan los principios de integración, de universalidad y de transformación.

---

<sup>9</sup> Los NDC son las siglas en inglés de contribuciones determinadas a nivel nacional (Nationally Determined Contributions).

### *Ilustración 5: Los 17 Objetivos de Desarrollo Sostenible (ODS)*



*Fuente: Alto Comisionado para la Agenda 2030*

Conviene aclarar que la aplicación y cumplimiento de estos objetivos y metas es responsabilidad de cada país, pudiendo prevalecer unos objetivos por encima de otros. Aun así, para alcanzar estas metas todo el mundo tiene que aportar, tanto el sector público como el sector privado y la sociedad en general. Por ello, cada país puede cambiar las formas, pero manteniendo el núcleo general (Gobierno de España, 2019). El nexo que hay entre las finanzas, las metas y objetivos es indudable y compele al sistema financiero a facilitar las necesidades globales contenidas en los ODS, para lo que deberá ser sujeto de una revolución y poder así llegar a la población desatendida, debiendo el sector privado enfocar esta responsabilidad como una oportunidad de conseguir beneficios tanto a corto como a largo plazo, para sí mismo y para la sociedad en su totalidad.

### **2.3. Otras iniciativas de desarrollo de finanzas sostenibles**

Queda claro que existe un compromiso global para potenciar las finanzas sostenibles con el objetivo de crear un mundo más responsable. Por ello, tanto el sector público como el sector privado trabajan para favorecerlo de forma paralela, asumiendo ambos un gran papel en su impulso y evolución, aunque por diferentes medios, que pasamos a desarrollar en los dos epígrafes siguientes (2.3.1. y 2.3.2.).

### **2.3.1. Iniciativas desde el sector público**

El sector público juega un papel fundamental en la aceleración de las finanzas hacia la sostenibilidad. Todas las instituciones públicas han llevado a cabo diversas acciones para favorecer los cambios que las instituciones privadas y demás colectivos deben asumir. Esta labor más estratégica y normativa se debe a que estas instituciones públicas tienen un horizonte temporal más amplio, alejándose del criterio de rentabilidad que predomina en el sector privado. El establecimiento de un marco de actuación para sentar unas bases a largo plazo que alineen intereses a cualquier nivel y la formulación de una normativa que exija inversiones y responsabilidades es tarea fundamental del sector público en todas sus variantes, debiendo también ejercer los oportunos seguimientos de su cumplimentación y el establecimiento de buenas prácticas.

A continuación, se presentarán diferentes iniciativas que se han formulado por parte de instituciones públicas, a saber: los Bonos Verdes; la creación del TCFD; la creación de la Red de Supervisores; el Plan de Acción de la Comisión Europea; los intentos de creación de una taxonomía; el Pacto Verde Europeo y, por último, el Fondo de Transición Justa.

Los Bonos Verdes, que surgieron en 2007 impulsados por el BEI, Banco Europeo de Inversiones<sup>10</sup>, se definen como bonos cuyos fondos se destinan exclusivamente a financiar proyectos verdes. Tienen como objetivo permitir y desarrollar el papel clave que los mercados de deuda pueden desempeñar en la financiación de proyectos que faciliten la sostenibilidad ambiental (International Capital Market Association, 2018). Los Bonos Verdes deben estar alineados con las cuatro condiciones contenidas en los Green Bond Principles (GBP<sup>11</sup>) y que se refieren a su uso, proceso de selección de proyectos, a la gestión de los fondos e informes preceptivos. Estos principios persiguen la integridad del mercado de Bonos Verdes a través de directrices sobre transparencia, publicidad y reporte de informes. Los GBP están formulados para su utilización por el mercado en su conjunto.

---

<sup>10</sup> Que es el órgano financiero comunitario de la Unión Europea.

<sup>11</sup> GBP hace alusión a los Green Bond Principles, que son los principios de los Bonos Verdes

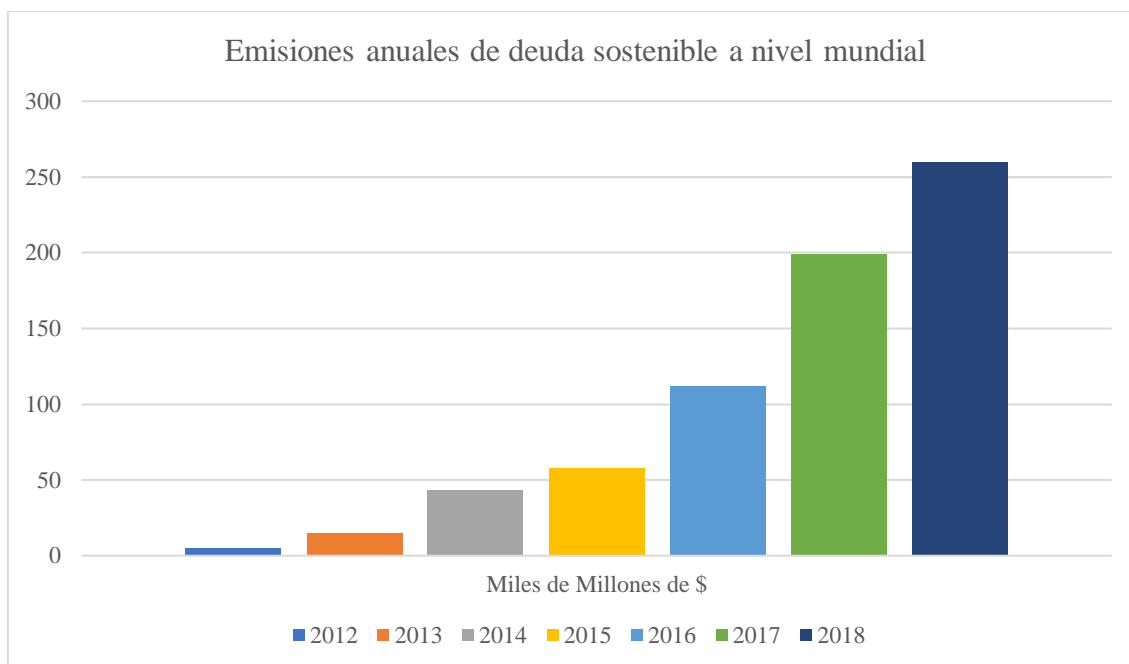
Proporcionan orientación a los emisores sobre los componentes clave involucrados en el lanzamiento de un Bono Verde creíble, ayudan a los inversores a promover la disponibilidad de la información necesaria para evaluar el impacto ambiental de sus inversiones en Bonos Verdes y a las entidades intermediarias a dirigir el mercado hacia nuevas prácticas que facilitarán futuras transacciones.

Los GBP exigen transparencia y claridad tanto en procedimiento como en la presentación de resultados, de modo que el resto de las partes interesadas del mercado entiendan las características de cualquier Bono Verde.

También exigen que los fondos deben estar destinados a inversiones en Proyectos Verdes como las energías renovables, la mejora de la eficiencia energética, el control de la contaminación, el transporte más limpio y la adaptación al cambio climático, entre otras. La gestión de los fondos por parte del emisor debe estar complementada por un auditor externo que certifique la correcta asignación de los fondos netos obtenidos por el Bono Verde. Esta acción de transparencia está acompañada de la creación por parte de los emisores de informes conteniendo datos actualizados sobre el uso de los fondos obtenidos, incluyendo la lista de los proyectos a los que se han asignado los fondos del Bono Verde, con una breve descripción y las cantidades asignadas a los mismos.

A pesar de que los Bonos Verdes salieron a la luz en 2007, no fue hasta 2013 cuando empezaron a ganar peso. A partir de ese momento su ascenso ha sido constante como queda reflejado en la Ilustración 6, cuyos datos son extraídos de Bloomberg New Energy Finance, que muestra en 2016 una inversión superior a los 87.000 millones de dólares en emisiones en todo el mundo, cifra que duplicaba los 42.000 millones de dólares invertidos de 2015, y esa cantidad no ha dejado de crecer hasta la actualidad llegando a más de 800.000 millones de dólares acumulados hasta 2019 según Bloomberg, siendo una de las claves su creciente innovación y diversificación. La cifra anual de bonos verdes emitida por empresas y organizaciones de todo el mundo acumulada a mayo del 2021 asciende a 148.900 millones de dólares (Climate Bonds Initiative, CBI).

**Ilustración 6: Emisiones anuales de deuda sostenible a nivel mundial, en miles de millones de \$ en el periodo 2012 a 2018.**



*Fuente: Elaboración propia con Excel, datos extraídos de Bloomberg*

El TCFD<sup>12</sup> (Task Force on Climate-related Financial Disclosures) fue otra iniciativa del sector público que surgió en 2015 en el marco del G20<sup>13</sup> por medio del FSB<sup>14</sup> (Consejo de Estabilidad Financiera) como respuesta a la preocupación creciente que había sobre los riesgos relacionados con el cambio climático. En aquel momento, y potenciado por el Acuerdo de París, existía la convicción de que los mercados financieros tendrían un papel fundamental en la lucha contra el cambio climático y sus consecuencias, y fue en ese contexto bajo el que se creó este Grupo. El TCFD elaboró en junio del 2017 una propuesta de reportes financieros relacionados con el cambio climático que pudiesen ser útiles para inversores, financiadores y aseguradoras. Las recomendaciones del TCFD pretenden ayudar a las empresas a comprender qué es lo que los mercados financieros demandan en materia de divulgación sobre los riesgos del cambio climático, y animan a las empresas a ajustar su divulgación a las necesidades de los *stakeholders* (FSB, 2015).

<sup>12</sup> El TCFD se refiere a las siglas de Task Force on Climate-related Financial Disclosures, que es un grupo de trabajo que se esfuerza para que las empresas informen a sus grupos de interés acerca de los riesgos relacionados con el cambio climático (EY, 2020).

<sup>13</sup> El G20 es el grupo de los países más poderosos del mundo que representan el 85% de la economía mundial (BBC, s.f.)

<sup>14</sup> El FSB es el Consejo de Estabilidad Financiera, creador del TCFD.

El informe de recomendaciones TCFD estaba estructurado en cuatro áreas temáticas: Modelo de gobierno, Estrategia, Gestión de riesgos y Métricas. Repsol, empresa que será objeto de análisis, se adhirió al TCFD en el año 2018.

En diciembre de 2017 llegó otra iniciativa en materia sostenible en el ámbito público, en este caso desde los Bancos Centrales y la Red de Supervisores, y que se llamó Red de Supervisores y Bancos Centrales para Enverdecer al Sistema Financiero<sup>15</sup>. Se trata del único foro mundial que reúne a Bancos Centrales y supervisores para luchar contra los riesgos del cambio climático. Esta red cuenta con el apoyo de una gran cantidad de miembros y de entes observadores. Entre ellos están diferentes Bancos Centrales de todo el mundo, siendo los miembros del comité de seguimiento los siguientes: Banco de México, Banco de Inglaterra, Banco Alemania, Banco de Singapur, Banco de Países Bajos, Banco de Finlandia, Banco Francés, Banco de Marruecos, Banco China y el Banco de Pagos Internacionales<sup>16</sup> (NGFS, 2019).

En el primer informe publicado por la NGFS se ofrecen seis recomendaciones para que bancos centrales, supervisores, responsables de la política pública e instituciones financieras desempeñaran un mejor papel en el “enverdecimiento” del sistema financiero y en la administración de los riesgos relacionados con el cambio climático y el medio ambiente. Las recomendaciones no son vinculantes y exponen las mejores prácticas identificadas por sus miembros para facilitar el involucramiento del sector financiero con los objetivos del Acuerdo de París (NGFS, 2019).

Las seis recomendaciones son las siguientes:

- 1- Integrar los riesgos relacionados con el clima en el monitoreo de la estabilidad financiera y la micro-supervisión.
- 2- Integrar los factores de sostenibilidad en la administración de la cartera.
- 3- Cerrar las brechas de información.
- 4- Crear conciencia y capacidad intelectual, así como fomentar la asistencia técnica y el intercambio de conocimientos.

---

<sup>15</sup> Que será tratado como NGFS por ser sus siglas en inglés

<sup>16</sup> En inglés se conoce como el BIS (acrónimo de Bank for International Settlements)



- 5- Lograr una divulgación de información relacionada con el clima y el medio ambiente robusta y consistente a nivel internacional.
- 6- Apoyar el desarrollo de una taxonomía verde de las actividades económicas.

El NGFS define también 3 líneas de acción principales: la supervisión macro-prudencial, las macro-finanzas, y por último la integración de las finanzas verdes (NGFS, 2019).

Sin embargo, la mayor aportación al desarrollo de finanzas sostenibles en Europa hasta la fecha ha venido por parte de la Comisión Europea y se concreta en el Plan de Acción para vincular el sistema financiero al desarrollo sostenible. Nació en marzo de 2018 con el objetivo de fomentar una economía más ecológica y limpia (CNMV, 2018).

Desde el nacimiento de este plan, la cantidad de flujos de financiación hacia activos sostenibles y actividades menos contaminantes no ha dejado de aumentar. Este crecimiento, propulsado por diferentes acuerdos que han quedado presentados previamente, no es suficiente aún para satisfacer las necesidades de financiación para las diferentes economías, ya que se estima que solo en Europa son necesarios 260.000 millones de euros de inversión adicional al año para lograr los objetivos europeos en materia de clima y energía de 2030 (Comisión Europea, 2019).

Esta falta de financiación para satisfacer al completo las necesidades se achaca a la inexistencia de una definición global de lo que es un proyecto o inversión sostenible. La existencia de una definición comúnmente aceptada podría reducir algunos de los riesgos relacionados con la inversión “sostenible”, como por ejemplo el blanqueo ecológico (*greenwashing*) ejercido por algunos inversores. Dicha definición global reduciría riesgos al introducir claridad entre las partes interesadas, dando la oportunidad a la taxonomía de ser un potente catalizador de flujos financieros hacia una economía sostenible bajo un entorno regulatorio y de apoyo político adecuado (OCDE, 2020).

Es por ello que la creación de una taxonomía sostenible ha sido uno de los intentos más relevantes desde el sector público. El primero fue el realizado por la Climate Bonds Initiative<sup>17</sup> que precedió a los esfuerzos de la Comisión Europea pues se trata de una organización global creada en 2009 centrada en los inversores y que moviliza los mercados de capital de deuda para soluciones climáticas. Consciente de la necesidad de una taxonomía, esta organización presentó en 2013 la Taxonomía de CBI, que identifica los proyectos y activos necesarios para proporcionar una economía baja en carbono, y suministra los criterios de evaluación de las emisiones de gases de efecto invernadero (GEI) en coherencia con el objetivo de calentamiento global de 2 grados fijado por el Acuerdo de París. Se actualiza periódicamente en base a la ciencia climática más reciente, la aparición de nuevas tecnologías y los criterios por sectores bajo los estándares de CBI (CBI, 2019). La evaluación de las emisiones de GEI consiste en cuatro categorías en forma de semáforo dependiendo del grado de sostenibilidad y contribución al objetivo de límite del calentamiento global de 2 grados de la actividad analizada (compatible automáticamente- verde; compatible si cumple con el indicador de evaluación - naranja; no compatible - rojo; se requiere más trabajo - blanco).

Los Principios de Bonos Verdes (GBP<sup>18</sup>) constituyeron otro intento de creación de Taxonomía, insuficientes por las limitaciones que contienen, pero que contribuyeron de manera clara a la expansión de emisiones de Bonos Verdes en estos años.

La heterogeneidad de las definiciones de los proyectos sostenibles generó la necesidad de crear una Taxonomía de ámbito europeo o continental que evitara el arbitraje entre entornos regulatorios nacionales. De este modo se eliminarían los *gaps* normativos evitando las diferencias de permisividad legislativa que pudiera ser aprovechada por algún operador financiero en su exclusivo beneficio al etiquetar como sostenibles de forma interesada y engañosa sus productos financieros.

A continuación, se expondrán los puntos clave de la nueva Taxonomía de la Comisión Europea que será aplicada en toda la UE<sup>19</sup>.

---

<sup>17</sup> Que será referido también como CBI, que son sus siglas en inglés (Climate Bonds Initiative)

<sup>18</sup> GBP son las siglas en inglés de los Principios de Bonos Verdes (Green Bond Principles)

<sup>19</sup> UE son las siglas de Unión Europea

En primer lugar, se marcan las condiciones que debe cumplir una actividad para que pueda ser considerada como sostenible, a saber:

- a) Que contribuya de forma directa y sustancial, evitando el oportunismo inversor, a la consecución de alguno de los siguientes objetivos medioambientales: mitigación o adaptación al cambio climático; protección de los recursos hídricos y marinos; impulso de la economía circular; la prevención de la contaminación, y recuperación de la biodiversidad.
- b) Que permita mantener una visión conjunta y global de los objetivos medioambientales, de manera que no perjudiquen unos a otros este objetivo macro marcado, cumpliendo de este modo con los criterios técnicos de selección que fueron redactados por un comité de 35 expertos creado por la CE<sup>20</sup> en 2018.
- c) Que cumpla con las garantías sociales mínimas, alineándose con la OCDE<sup>21</sup> para Multinacionales, con los principios de las Naciones Unidas sobre las empresas y los derechos humanos en el trabajo, y con la normativa de la UE en los ámbitos de medio ambiente, salud y seguridad, y sostenibilidad social.

Es fundamental destacar que la Taxonomía se centra en actividades y no en organizaciones, de modo que una organización puede llevar a cabo actividades económicas sostenibles y no sostenibles al mismo tiempo.

Otra de las claves del Reglamento de la Taxonomía de la CE es la divulgación de información medioambiental, de manera que refuerce las normativas relacionadas (el Reglamento UE 2019/2088 y la NFRD<sup>22</sup>). De esta manera, el Reglamento amplía las obligaciones de transparencia sobre la información no financiera de las empresas que sobrepasen los 500 empleados.

---

<sup>20</sup> CE son las siglas de la Comisión Europea

<sup>21</sup> OCDE son las siglas de la Organización para la Cooperación y el Desarrollo Económicos

<sup>22</sup> NFRD son las siglas de Non-Financial Reporting Directive

Antes de exponer los criterios técnicos de la Taxonomía, es preciso definir a que hacen referencia los términos de mitigación y adaptación. Las medidas de mitigación se ponen en marcha para evitar el aumento de las emisiones contaminantes (como el proceso de mantener el incremento de la temperatura media mundial por debajo de los 2 °C y de continuar limitarlo a 1,5 °C respecto de los niveles preindustriales), mientras que las medidas de adaptación son acciones para reducir la vulnerabilidad ante las consecuencias del cambio climático.

Los criterios técnicos referentes al objetivo de mitigación son fundamentalmente cuantitativos, siendo diferentes en función del tipo de actividad económica que puede contribuir al objetivo de mitigación, distinguiendo el reglamento tres categorías diferentes:

- a) Actividades bajas en carbono que no producen emisiones, lo hacen en niveles muy bajos o que utilizan tecnologías de secuestro de carbono.
- b) Actividades de transición que no tienen mejor alternativa en términos de sostenibilidad y que, además, apoyan la transición.
- c) Actividades facilitadoras que hacen posible que otras actividades puedan realizar una contribución sustancial al objetivo de mitigación y que no dañan otros objetivos medioambientales

En cuanto al objetivo de adaptación, los criterios técnicos no son cuantitativos, ya que dependen del contexto o localización de la actividad. Las dos clases que diferencia el TEG<sup>23</sup> son:

- a) Actividades adaptadas. Se trata de aquellas actividades que están preparadas para no ser vulnerables.
- b) Actividades facilitadoras. Son aquellas que facilitan la adaptación de otras actividades y que están preparadas para los riesgos por sí mismas.

---

<sup>23</sup> El TEG se refiere al Grupo técnico de expertos de la UE en Finanzas Sostenibles, asesor de la UE para la creación de la Taxonomía

Tras las aportaciones mencionadas, el TEG fue sustituido por la Plataforma de Finanzas Sostenibles, formada por 50 miembros además de los 7 de elección directa y los 9 observadores, donde participan países fuera de Europa como Argentina, Chile, China, India, Kenia, Marruecos o Canadá.

La Taxonomía, sin embargo, no está terminada, ya que sigue en formación y se espera que siga evolucionando hasta convertirse en referencia en un futuro próximo.

Tras los comentados intentos de creación de una taxonomía, conviene presentar el plan que estableció el Pacto Verde Europeo en el 2019, que establece el objetivo de lograr la neutralidad climática en la UE en el 2050 y avanzar en el cumplimiento de los ODS. Para ello es necesario impulsar el uso eficiente de los recursos definiendo una transición energética, la implantación progresiva de la economía circular y la restauración de la biodiversidad. Queda pendiente convertir este Pacto en una obligación legal, de manera que se asegurara mejor su cumplimiento.

Poco tiempo después del Pacto Verde Europeo, a principios de 2020, la UE presentó el Mecanismo de Transición Justa, que forma parte del plan de recuperación llamado NextGenerationEU, que será comentado de forma más completa en el epígrafe 2.5.1. Se trata del apoyo financiero y técnico a las personas, las empresas y las regiones más afectadas para la consecución de una transición sostenible (UE, 2020).

### **2.3.2. Iniciativas desde el sector privado**

Como se introdujo previamente, en la evolución de las Finanzas Sostenibles han tenido gran responsabilidad tanto el sector público como el privado. Se ha expuesto como el sector público ha tenido un papel más normativo en el sentido de haber marcado ciertos límites, haber planificado los esfuerzos futuros a realizar, haber definido estrategias y haber estimulado al resto de las partes implicadas en el tema.

Una vez definidos los esfuerzos públicos, es preciso repasar algunos de los esfuerzos realizados desde instituciones privadas para impulsar las Finanzas Sostenibles, exponiendo en primer lugar las motivaciones y, por último, los distintos instrumentos.

Las motivaciones más destacables de las iniciativas privadas son:

- Incrementar la inversión sostenible.
- Realizar de obras de anticipación a posibles restricciones o limitaciones desde el sector público.
- Las presiones recibidas por parte de los inversores.
- Integrar la sostenibilidad en un plan estratégico.

Por tanto, podemos ver que la iniciativa privada está propulsada principalmente por dos motores: la presión pública y la identificación de riesgos. Es importante resaltar que todos los esfuerzos no son producto de labores altruistas o filantrópicas en busca del bien común, sino que también responden a otras cuestiones que pueden alterar la viabilidad y la estabilidad financiera, y que han potenciado la preocupación creciente del sector financiero privado. Una vez se ha hecho evidente la importancia del tema, el sector privado ha tenido que analizar los riesgos y las oportunidades que pueden surgir en este contexto marcado por un cambio del modelo económico hacia uno más sostenible.

Comenzando por los riesgos, el más visible e importante es el que se refiere a la futura obsolescencia de multitud de activos que son propiedad de instituciones privadas. Una obsolescencia o pérdida de valor que vendrá como consecuencia del cambio de modelo económico, que significará el final de vida útil de gran cantidad de activos. Esta pérdida puede llegar a los 1,6 billones de dólares en activos del sector petrolero, del gas y del carbón (The Carbon Tracker Initiative, 2018). Este principal riesgo es el que desencadena una de las mayores oportunidades que surgen para el sector privado y que es la necesidad de inversión mundial en infraestructuras. Oportunidades asociadas al cambio de modelo económico, que tendrá al sector energético como uno de los principales, y que, según la Comisión Europea, necesitará unas inversiones de más de medio billón de dólares al año.

Además de la obsolescencia de infraestructuras hay tres factores más que han hecho que inversores institucionales y gestores de activos tengan una mayor implicación en la “revolución sostenible”. El primero de ellos es la mayor rentabilidad que aportan aquellas inversiones que adoptan criterios sostenibles, como han demostrado diferentes estudios. Otro factor es el fortalecimiento de la gestión del riesgo por la creciente preocupación por parte de empresas privadas de la gran repercusión que pueden tener en la cuenta de

resultados noticias negativas como la insatisfacción de sus empleados, problemas con el tratamiento de residuos o diversos accidentes con daños medioambientales. Por último, el alineamiento con los *stakeholders* como consecuencia del creciente interés en las preocupaciones de sus principales grupos de interés. Todos estos factores están presididos por un objetivo general de toda inversión por parte de una organización privada que es maximizar la rentabilidad ajustando el riesgo.

En cuanto a los instrumentos, se van a desarrollar los siguientes: Bonos Verdes; bonos de transición, préstamos verdes; financiación de proyectos; cédulas y titulaciones; derivados financieros; Principios de Banca Responsable; y, por último, los Principios para la sostenibilidad en los seguros.

El primer instrumento mencionado son los Bonos Verdes, y que ya hemos expuesto en las iniciativas públicas (apartado 2.3.1.) pero que en realidad se trata de una iniciativa público-privada, por lo que nos centraremos en otros instrumentos de financiación sostenible en el ámbito del sector privado.

Existen otros bonos que se denominan *bonos de transición*, que están destinados a financiar empresas que actualmente no cumplen la totalidad de requisitos planteados por la taxonomía para denominarse Bonos Verdes pero que tienen un papel fundamental en la consecución de una economía más sostenible.

Paralelamente al crecimiento de los bonos sostenibles, han aparecido los préstamos verdes como herramienta de financiación en expansión. Al igual que los bonos, los préstamos verdes también cuentan con unos principios llamados Green Loan Principles, de ICMA<sup>24</sup>, que detallan la definición del compromiso verde, el proceso de evaluación de los proyectos financiados, la gestión de los fondos y la labor de *reporting*. Según datos de Bloomberg Finance, los préstamos verdes han pasado de movilizar alrededor de 15 mil millones de euros en 2012 a superar los 50 mil millones de euros en 2018. España destaca como uno de los principales emisores de préstamos verdes (Banco de España, 2019).

---

<sup>24</sup> ICMA son las siglas de la Asociación Internacional de Mercados de Capitales

Existe también otro vehículo inversor utilizado tanto por el sector público como el privado que es la *financiación de proyectos*, que disfrutan de descuentos en el coste financiero tras la evaluación de la sostenibilidad del proyecto.

Otros instrumentos que provienen principalmente de la iniciativa privada son las *cédulas* y las *titulizaciones*. En primer lugar, las cédulas verdes destacan por el compromiso que adopta el emisor en usar los fondos captados en proyectos sostenibles. Además de tener la masa patrimonial como garantía, algunos emisores cuentan a su vez con una cantidad de activos verdes como respaldo de la cédula. En cuanto a las titulizaciones, las entidades emisoras, mediante dicho instrumento, sacan de su balance determinados activos, financiando su enajenación mediante la venta de un instrumento denominado *bono de titulización*. Estos bonos se diferencian de las cédulas en el respaldo que hay detrás de las inversiones, siendo en el caso de los bonos de titulización los activos enajenados, mientras que en las cédulas el respaldo es la masa patrimonial de la entidad (Banco de España, 2019).

También existen los *derivados financieros sostenibles* que, aunque aún no hay gran oferta, constituyen una herramienta de cobertura de riesgos climáticos por medio de los derivados del clima o los derivados energéticos.

El sector bancario, además del involucramiento en diferentes instrumentos ya comentados, también ha aumentado su compromiso en la fijación de objetivos y de marcos de actuación a seguir en el largo plazo. Como respuesta a ello, en 2019, con el apoyo de más de 130 bancos, se lanzaron los Principios de Banca Responsable, que conformaron un apoyo a las diferentes iniciativas tanto públicas como privadas (UNEP-FI, 2012). Uno de los puntos principales de estos principios es la transparencia y la mayor alineación de su actividad con la necesaria revolución sostenible.

El sector asegurador es uno de los más afectados por el riesgo, ya que está expuesto de forma directa al modelo de crecimiento no sostenible. Por ello, el sector privado asegurador ha tenido una gran participación en la gestión sostenible. Una contribución de este sector son los Principios para la Sostenibilidad en los Seguros, que fueron aprobados en el año 2012 y que ya cuentan con más de 70 entes vinculadas. Consideran la sostenibilidad desde un enfoque ambiental y desde otro social debido a la volatilidad e



incertidumbre de la demanda de aseguradoras por la frecuencia y severidad de algunos desastres naturales, y cuya evolución se vio en la Ilustración 3.

## **2.4. Criterios ASG**

Los criterios ASG (ESG<sup>25</sup> en inglés) constituyen un conjunto de recomendaciones a seguir en la toma de decisiones empresariales sobre inversiones en activos, mediante el análisis de información fundamentalmente extra financiera que mejora la gestión del riesgo, midiendo la exposición de una empresa a cuestiones no financieras. Los criterios que utiliza son ambientales, sociales y de gobernanza en función: del impacto que tiene la actividad de una empresa en el medioambiente; de la relación de la empresa con la sociedad, especialmente con sus *stakeholders*; de la forma en la que la empresa es administrada y del cumplimiento por parte de la empresa de diferentes reglamentos. En el epígrafe 5.3 del presente trabajo se estudiará a la empresa Repsol en cuanto a los criterios ESG.

## **2.5. La pandemia y la sostenibilidad**

Nadie ni nada es ajeno a la pandemia que vivimos desde principios del año 2020. Un año después, nos encontramos inmersos en una crisis sanitaria y económica que se ha llevado por delante miles de vidas y empleos, entre otras cosas. Ningún país ha salido indemne ante esta situación. A pesar de ello, las Finanzas Sostenibles parecen haber resistido y el COVID-19 ha acelerado la “revolución sostenible”, algo que puede intuirse en el informe presentado por Deloitte, llamado “2020 Global Impact Report”, que proyecta que en 2024 más del 50% de los activos gestionados en Estados Unidos estarán alineados con la sostenibilidad<sup>26</sup>. Esta afirmación puede respaldarse incluso en Europa con datos, como la comparación entre los rendimientos interanuales de dos índices bursátiles “hermanos” como son el MSCI Europe ESG y el MSCI Europe, que muestra una mejora en rendimiento del 4% para el índice que incluye los ESG.

---

<sup>25</sup>ESG (Environmental, Social, Governance) son las siglas en inglés de los ASG (ambiental, social y gobernanza)

<sup>26</sup> En 2024 más de la mitad de los activos en Estados Unidos estarán alineados con la sostenibilidad (Deloitte, 2020).

Morningstar, por su parte, apunta que los fondos sostenibles han crecido, mientras que otros fondos de inversión han sufrido enormes pérdidas en el último año<sup>27</sup>.

En cuanto la renta fija, un informe del Multidisciplinary Digital Publishing Institute con el nombre “What future of Green Bond Market” expone que solo en Europa, el mercado de bonos verdes se duplicará en 2021<sup>28</sup>.

Detrás de estos datos puede haber diferentes explicaciones. Entre ellas, podríamos afirmar que gran parte de los inversores en activos sostenibles invierten en ellos por su visión a largo plazo que no se ha detenido con la pandemia. Además, muchos fondos de inversión usan criterios ESG para sus decisiones de inversión desde antes de la pandemia, algo que tampoco se ha frenado.

Otra explicación es que muchas instituciones públicas, en sus planes de recuperación por la pandemia, han articulado sus esfuerzos en las finanzas sostenibles poniendo el foco en el incentivo a las tecnologías digitales, lo cual puede llevar a que las finanzas sostenibles sean la tendencia número uno entre las modalidades de inversión en los próximos años.

Por otro lado, la pandemia le ha dado al planeta una tregua en la guerra que mantiene el crecimiento económico con la sostenibilidad, de manera que la reducción de la actividad económica ha reducido en 2020 las emisiones de los GEI, siendo esta una de las pocas buenas noticias que esta terrible pandemia ha ocasionado.

Dentro de este contexto, cabe presentar el Fondo de Recuperación Europeo, el NextGenerationEU, que será desarrollado en el próximo epígrafe 2.5.1.

---

<sup>27</sup> Los fondos sostenibles han experimentado un crecimiento en torno al 15%, mientras que otros fondos de inversión han sufrido enormes pérdidas en el último año (Morningstar, 2021)

<sup>28</sup> En Europa, el mercado de bonos verdes se duplicará en 2021 (MDPI, 2020)

### **2.5.1. NextGenerationEU**

El NextGenerationEU es un plan de recuperación dotado con 750.000 millones de euros que contribuirá a reparar los daños económicos y sociales causados por la pandemia de coronavirus, que tiene el objetivo de que la Europa posterior a la COVID-19 sea más sostenible, más digital, más resiliente y mejor adaptada a los retos actuales y futuros. El Plan está desglosado en diferentes fondos y supone la iniciativa de mayor envergadura económica jamás realizada por Europa. Destaca el Mecanismo Europeo de Recuperación y Resiliencia, que supone la base del Plan de recuperación y que aportará 672.500 millones de euros (360.000 millones de euros en préstamos y 312.500 millones de euros en subvenciones) disponibles para apoyar las reformas e inversiones de los países de la UE (Plan de recuperación para Europa, 2020). Además de este, hay otros fondos y programas dentro del plan de recuperación como la Ayuda a la Recuperación para la Cohesión y los Territorios de Europa (REACT-UE), Horizonte 2020, InvestEU, Desarrollo Rural o el Fondo de Transición Justa (FTJ<sup>29</sup>). Entre ellos, el FTJ guarda gran relación con las Finanzas Sostenibles ya que busca que la transición hacia una economía climáticamente neutra se haga de forma equitativa mediante la inversión de más de 150.000 millones de euros durante el período 2021-2027 en las regiones más afectadas para mitigar el impacto socioeconómico de la transición, que ya fue introducido previamente (Comisión Europea, 2020b).

---

<sup>29</sup>FTJ hace referencia al Fondo de Transición Justa

### 3. CONCEPTOS BÁSICOS DE TEXT MINING, WEBSCRAPING Y ANÁLISIS DE SENTIMIENTOS

Antes de comenzar con el estudio y planteamiento de la metodología utilizada en el mismo, es preciso definir qué es el *Text Mining*, el *Webscraping* y el análisis de sentimientos.

Se estima, según Oracle, que el 90% del total de los datos son no estructurados<sup>30</sup>. Por tanto, únicamente un 10% es estructurado, lo que quiere decir que una décima parte de los datos están contenidos en una base o algún otro tipo de estructura de datos (Rayón, 2020b). Para tratar con la gran mayoría de ellos, el *Text Mining* se presenta como una gran opción. El *Text Mining*, también llamado minería de textos, es el uso de técnicas computacionales para extraer información y conocimiento de alta calidad del texto (Rayón, 2020b).

A pesar de que se suele confundir con la recuperación de la información (*Information Retrieval*) (Hearst, 1999), que consiste en la recuperación automática de documentos relevantes, el *Text Mining* se refiere a examinar una colección de documentos y descubrir información no contenida en ningún documento individual (Nasukawa, 2001). Por tanto, podemos concluir que se encarga de descubrir información que no existía previamente de forma explícita. Estas técnicas de minería de textos tienen tres posibles fases principales.

En primer lugar, la fase de procesamiento léxico y sintáctico, que se refiere a: reconocer los límites de las palabras, normalizar el texto (segmentar palabras del texto, normalizar sus formatos y segmentar las frases del texto), la tokenización (o segmentación), la lematización y el *stemming* por parte del análisis léxico. Por parte del análisis sintáctico encontramos: *parsing* para determinar la estructura de una frase y etiquetar cada palabra con su categoría gramatical.

---

<sup>30</sup> El 90% del total de los datos son no estructurados (Vicenteño, 2017)

En segundo lugar, la fase de procesamiento semántico que se refiere a localizar y clasificar palabras del texto que hacen referencia a elementos concretos como nombres de personas, lugares o entidades para poder encontrar patrones, tendencias, relaciones entre conceptos en el texto.

Por último, la fase de extracción de características extra-semánticas, de modo que se pueda realizar un análisis de sentimiento (Rayón, 2020b).

Una vez definido lo que es el *Text Mining*, se definirá el *Webscraping*, ya que es una técnica que se utilizará tanto para el estudio de las menciones en el New York Times como para el de la información sacada de Twitter.

El *Webscraping* es una técnica para transformar la información de sitios web y redes sociales que se encuentran de forma desestructurada y transformarla en datos estructurados de forma automatizada (Grupolambda, 2021).

Por último, el análisis de sentimientos, que se refiere al proceso mediante el cual se determina el aspecto emocional que tienen una serie de términos, siendo especialmente útil en la monitorización de las redes sociales, y es por ello por lo que se va a hacer uso de esta técnica para extraer los sentimientos y emociones de una serie de *tweets*, siendo Twitter la red social escogida por el gran volumen de datos que puede proporcionar.

## 4. METODOLOGÍA

En este trabajo se han adoptado diferentes metodologías que pertenecen al *Business Analytics*, que serán presentadas y explicadas a lo largo de este capítulo. Para ello, se ha dividido el capítulo en:

- Descarga de datos
- Análisis exploratorio y de sentimiento

### 4.1. Descarga de datos

#### 4.1.1. Descarga de datos del New York Times

En primer lugar, se ha considerado preciso ver la evolución de diferentes términos relacionados con el tema en varios intervalos temporales para poder visualizar el interés que han ido despertando. Tras la elección del medio de comunicación se ha observado la aparición de diferente terminología sostenible en artículos de negocios.

Se ha escogido el New York Times para este análisis por diferentes razones. Por un lado, por el acceso gratuito que se nos brinda el propio NYT a su base de datos. Por otro lado, el New York Times cuenta con gran volumen de artículos mensuales de hechos globales de amplia tipología. En este trabajo se va a centrar la búsqueda en la sección *Business* al ser la que mejor encaja con el objetivo.

Para ello, se ha tenido que crear una clave de *developer* en el New York Times para poder tener acceso a su base de datos. También ha sido necesaria la creación de una *APP*, para después poder obtener una *API Key* que permitiera la autenticación en cada llamada que se hizo a la API<sup>31</sup> del NYT. Una vez obtenido lo anterior, se ha procedido a la creación de un *script* en el software de programación *RStudio* para el tratamiento de los datos.

---

<sup>31</sup>API son las siglas en inglés de Application Programming Interface, que se trata de un conjunto de peticiones http estructuradas que devuelven datos en un formato ágil y eficiente (Rayón, 2020).

Para cumplir el objetivo de visualizar la evolución en cuanto a menciones en artículos en el NYT de la terminología sostenible, se han tenido que ejecutar distintas funciones que serán presentadas en el epígrafe 4.2.1.

#### **4.1.2. Descarga de datos de las cotizaciones de Repsol.**

En primera instancia se descargaron de *Yahoo Finance* los valores de la cotización de Repsol en el MCE<sup>32</sup> desde enero de 2015 (año en el que se adoptaron ciertos acuerdos en materia sostenible como el Acuerdo de Paris o la Agenda 2030 y los ODS), hasta el 31 de diciembre de 2020.

Para ello, se utilizó el programa *RStudio*, mediante el paquete *quantmod*, que está diseñado para asistir a los operadores cuantitativos en el desarrollo, pruebas y despliegue de modelos de trading basados en las estadísticas (Tellechea, 2017). Tras la descarga del paquete, se usó una función llamada *getSymbols*, gracias a la cual se han podido descargar de forma inmediata todas las cotizaciones semanales de Repsol desde el año 2015 hasta 2020. En la Ilustración 7 se puede ver un extracto del código utilizado.

Todos los datos descargados son semanales y están ordenados según su fecha. Su valor está representado en euros y en formato de distintas columnas: *Date*, *Open*, *High*, *Low*, *Close*, *AdjClose* y *volumen*. El precio de cierre (*Close*) es el precio de cierre después de los ajustes por *splits*. El precio de cierre ajustado (*AdjClose*) es el precio de cierre ajustado por dividendos y *splits*. La decisión de escoger la frecuencia semanal de datos se debe a que una elección mensual podría haber provocado una pérdida de información relevante para la elaboración del epígrafe 5.4. en el que se estudia la correlación entre la cotización de Repsol y el sentimiento promedio extraído de los *tweets* y que requiere intervalos temporales de captura de datos idénticos. Por otro lado, se desechó la frecuencia diaria tras considerar que la escasez de *tweets* habiendo establecido una frecuencia de datos diaria en intervalo de muestra de 5 años completos podría añadir complicaciones adicionales de interpretación de correlación de datos.

---

<sup>32</sup>El MCE es el Mercado Continuo Español, que conecta las cuatro bolsas españolas (Madrid, Barcelona, Valencia y Bilbao). Se ha utilizado el valor que aporta el MCE al ser el que posibilita que la acción de Repsol cotice simultáneamente en las distintas bolsas españolas, siendo la empresa de origen español.

*Ilustración 7: Extracto del código usado en RStudio para la descarga de las cotizaciones de Repsol en el MCE*

```
install.packages("quantmod")  
library(quantmod)  
getSymbols("REP.MC",src="yahoo")  
head(cotizaciones)  
tail(cotizaciones)
```

*Fuente: Elaboración propia en RStudio.*

### **4.1.3. Descarga de datos de Twitter.**

Se ha decidido utilizar la red social Twitter para analizar los sentimientos de unos *tweets* determinados y su posible impacto en la cotización de Repsol ya que se considera que un solo tweet puede conllevar grandes repercusiones en diferentes variables económicas, no siendo la cotización bursátil una excepción. Un ejemplo actual de lo anterior son los grandes cambios en el mercado de criptomonedas que lleva ocasionando Elon Musk mediante su cuenta de Twitter<sup>33</sup>.

Tras sopesar diferentes opciones para la descarga de tweets, entre las que destaca la prueba con la librería *GetOldTweets3* de *Python*, que no fue utilizada al considerarse obsoleta, se ha decidido utilizar la librería *snsrape* de *Python*, la cual ha permitido descargar *tweets* mediante la especificación de *keywords*<sup>34</sup>, fecha de inicio de búsqueda, fecha final de búsqueda e idioma. En *RStudio* no se consiguió realizar la descarga de *tweets* en el período seleccionado, dejando únicamente descargar *tweets* de forma semanal y con un límite de volumen de descarga inferior al que nos impone el *script* usado de *Python*, siendo esta la razón por la cual se utilizó el software de *Python* para la descarga.

En la siguiente Ilustración 8 se puede observar cómo se realizó la descarga en *Python*.

---

<sup>33</sup>Elon Musk es un físico, emprendedor, inventor y magnate sudafricano que está teniendo un gran impacto en diferentes criptos como Doge Coin o Bitcoin a través de *tweets*

<sup>34</sup>El término *keywords* se refiere a las palabras clave



### *Ilustración 8: Extracto del código usado en Python para la descarga de tweets*

```
from platform import python_version
print(python_version())
!pip install sncrape
#!pip install vaderSentiment
import pandas as pd
import numpy as np
import sncrape.modules.twitter as sntwitter
import csv
import matplotlib.pyplot as plt

# Set maximum tweets to pull
maxTweets = 20000
# Set what keywords you want your twitter scraper to pull
keyword = 'Repsol Sostenibilidad'
#Open/create a file to append data to
csvFile = open('pfe_tweets_resultPROSI.csv', 'a', newline='', encoding='utf8')
#Use csv writer
csvWriter = csv.writer(csvFile)
csvWriter.writerow(['id', 'date', 'tweet',])

# Write tweets into the csv file
for i, tweet in enumerate(sntwitter.TwitterSearchScrapper(keyword + ' lang:es since:2015-01-01 until:2020-12-30
                        if i > maxTweets :
                            break
                        csvWriter.writerow([tweet.id, tweet.date, tweet.content])
csvFile.close()
```

*Fuente: Elaboración propia en Python*

Como se puede observar en la Ilustración 8, las *keywords* empleadas fueron “Repsol” y “Sostenibilidad”. En primer lugar, la elección de “Repsol” se debe a que es el nombre de la compañía objeto de nuestro estudio. En segundo lugar, “Sostenibilidad”, debiéndose esta elección a que, tras probar diferente terminología sostenible, la unión de “Repsol” y “Sostenibilidad” fue la que mayor cantidad y calidad de tweets nos brindaba. Algunos de los términos “sostenibles” que se probaron fueron: “carbón”, “carbono”, “sostenible”, “Agenda 2030”, “emisiones” o “renovable” pero, como ya hemos explicado, ofrecían un número de tweets menor a 100, siendo su calidad similar o menor al de las *keywords* finalmente seleccionadas.

En cuanto a la elección de fechas, se estipuló que la de inicio fuese el 1 de enero de 2015 y la final el 31 de diciembre de 2020. La decisión de establecer dicha fecha de inicio se debe a diferentes razones entre las que destaca la proximidad de esta fecha a la aparición de diferentes planes de acción en materia sostenible, que permite la obtención de mayor cantidad de *tweets*. La elección de fecha final se debe a que se ha considerado como la mejor opción por incluir 5 años completos y ser una fecha actual.

Por último, se ha programado que se realice la búsqueda de los *tweets* en español ya que no existían *tweets* en inglés usando las *keywords* “Repsol” y “sustainability”, lo que puede tener lógica al ser una empresa española y que, aunque esté presente en 31 países

diferentes, el mayor volumen se concentra en países hispano hablantes donde dispone tanto de actividades de refino como de distribución y venta al por menor, siendo la participación en países no hispano hablantes en actividades extractivas de petróleo a través de participaciones financieras minoritarias que no habilitan para ser operadores de dichos campos petrolíferos, y por lo tanto con poca presencia mediática.

El número total de *tweets* obtenidos en la descarga es de 334, número que es muy reducido en comparación con los millones de *tweets* que se generan al día, pero que puede tener su explicación, entre otras posibles razones, en que algunos usuarios hayan borrado los dichos *tweets* o sus cuentas a lo largo de estos 5 años, reduciendo notablemente con ello el número de *tweets* disponibles.

Los datos descargados conforman un *dataset* que incluye distintas columnas: el Id del usuario, la fecha de envío de cada *tweet* y el contenido del *tweet* en cuestión (columna: *tweet*).

## **4.2. Análisis exploratorio y de sentimiento**

### **4.2.1. Menciones en artículos del New York Times**

En la descarga de datos se describió como se actuó para tener acceso a los datos de la *API* del New York Times. En este epígrafe se presentará como se ha procedido para obtener las evoluciones de los diferentes términos sostenibles.

Los paquetes necesarios para su elaboración son el *httr*, que se descarga para hacer llamadas a una *API*, y el *jsonlite*, que sirve para decodificar el resultado que viene en un JSON<sup>35</sup> a un *dataframe*.

---

<sup>35</sup>JSON son las siglas de JavaScript Object Notation

Una vez añadida la *Apikey* (se puede ver en la Ilustración 9), se ha incluido la base URL<sup>36</sup> sobre la que lanzar las consultas para los datos necesarios, además del *querying*, que recoge los documentos que coinciden con un criterio dado (en este caso será la terminología sostenible) y los clasifica (Rayón, 2020a). Ambas acciones pueden verse en la Ilustración 10.

**Ilustración 9: Llamada a la Api key**

```
apikey <- 'k9lukrGJJYPoDhuHTLuF13jcHB1ahRxd'
```

Fuente: Elaboración propia en Rstudio

**Ilustración 10: Funciones para la base URL y el querying**

```
base_url <- "https://api.nytimes.com/svc/search/v2/articlesearch.json"  
r <- GET(base_url,  
         query=list(q="Green Bonds", "api-key"=apikey))  
r
```

Fuente: Elaboración propia en Rstudio

Con estas acciones se obtiene un código que nos devuelve la *API*. Tras esto, se ha cargado la librería *jsonlite* para decodificar la respuesta y así poder crear un *dataframe* a partir de la respuesta.

Una vez se han realizado las acciones previas, se crean las funciones necesarias para poder obtener diferentes gráficos (*ggplots*) para cumplir el objetivo fijado. Para ello, en primer lugar, se ha tenido que fijar el intervalo temporal para cada búsqueda (ejemplo del procedimiento en la Ilustración 11), variando según el término en cuestión, ya que carece de sentido que compartan todos el mismo por tener fechas de creación distintas.

---

<sup>36</sup> URL son las siglas de Uniform Resource Locator

### ***Ilustración 11: Funciones para establecer intervalos temporales***

```
r <- GET(base_url, query=list(q="Green Bonds",
                             fq="section_name:(Business)",
                             "api-key"=apikey,
                             "begin_date"=20060101,
                             "end_date"=20210430))

r
respuesta <- fromJSON(content(r,"text"))
respuesta$response$meta

json <- content(r, 'parsed')
json
json$status
json$response$meta
```

*Fuente: Elaboración propia en Rstudio*

Una vez se ha establecido el intervalo de tiempo (fecha de inicio y fecha de final), se ha incluido una función para contar las apariciones en búsquedas, de manera que posteriormente pueda obtenerse su evolución. Se puede ver un extracto de la función utilizada para su obtención en la Ilustración 12.

### ***Ilustración 12: Funciones para contar las apariciones de un término en un intervalo de tiempo determinado.***

```
nyt_count <- function(q, date1, date2){
  r <- GET(base_url, query=list(q=q,
                               "api-key"=apikey,
                               "begin_date"=date1,
                               "end_date"=date2))

  json <- content(r, "parsed")
  return(json$response$meta$hits)
}
```

*Fuente: Elaboración propia en Rstudio*

Una vez llegados a este punto y habiendo establecido todo lo anterior, se está en disposición de llamar a la función anterior a contar apariciones en un periodo dado y, con ello, obtener los gráficos a partir de la función *plot*. Se puede ver en Ilustración 13.

***Ilustración 13: Funciones para contar las apariciones de un término en un intervalo de tiempo determinado y para crear un gráfico con su evolución.***

```
counts <- nyt_years_count(q="Green Bonds", yearinit=2006, yearend=2021)
plot(2006:2021, counts, type="l", main="Menciones de Green Bonds por año en NYT",
     xlab="Año", ylab="Número de artículos")
```

*Fuente: Elaboración propia en Rstudio*

Con esta serie de funciones, por tanto, se puede obtener la evolución de diferentes términos que serán interpretados posteriormente en el análisis de resultados (capítulo 5).

#### **4.2.2. Análisis de sentimiento**

Para el análisis de sentimiento ha sido necesario realizar un análisis léxico y un análisis semántico. Para ambos, se ha utilizado el software de programación *RStudio* que mediante los diferentes paquetes que incluye, resulta muy adecuado para dichos análisis.

- Análisis léxico

En primer lugar, se ha llevado a cabo un procesamiento léxico y sintáctico. Se ha utilizado el paquete *syuzhet*, que fue desarrollado por Matthew Jockers en 2015, y que trabaja con cuatro diccionarios de sentimientos: *Bing*, *Afinn*, *Stanford* y *NRC*. Este paquete permite transformar los *tweets* en un vector de palabras para su análisis de forma individual.

Se han reconocido los términos (*tokens*) con la función *get\_tokens()*, lo que nos ha permitido deshacernos de la puntuación en el texto y sacar una lista de un total de 11.344 palabras en nuestro caso.

A su vez, se ha realizado un *script* en *RStudio* diferente para sacar otras herramientas que sirvan de apoyo para extraer conclusiones. En él, se ha normalizado el texto, tras separar las palabras, eliminando las *stopwords*<sup>37</sup> con el paquete *stopwords* (*tokenización*).

---

<sup>37</sup>Las *stopwords* son palabras de parada, que se consideran irrelevantes al aparecer con gran frecuencia

Se ha creado un *dataframe* con diferentes términos eliminados, ya que se repetían con asiduidad:

"Repsol", "sostenibilidad", "sostenible", "Director", "@repsol", "luis", "dueñas", "cabra", "general", "#repsol, "@fundacionrepsol" o "#sostenibilidad" ", además de preposiciones. Tras ello, se ha realizado la labor de *stemming*, que consiste en reducir los términos de los tweets a sus raíces.

- Análisis semántico

En nuestro análisis semántico hemos utilizado el diccionario *NRC* (desarrollado por Saif M. Mohammad), ya que cuenta con el idioma español, con valores de sentimiento negativo o positivo y con ocho emociones (enfado, anticipación, disgusto, miedo, alegría, tristeza, sorpresa y confianza). Antes del uso de este paquete, se han debido tener en cuenta diferentes consideraciones: el léxico en español es una traducción directa realizada por traducción automática, que, gracias a ser el idioma español, cuenta con gran fiabilidad a diferencia de otros idiomas con los que trabaja el paquete; la valencia sentimental y emocional tiene un sesgo cultural y temporal; y, por último, este paquete no interpretará la negación de un sentimiento positivo (Isasi, 2021).

Una vez obtenido el vector de *tokens*, para el análisis se ha utilizado la función *get\_nrc\_sentiment* que ha proporcionado los sentimientos del texto. Para ello, se ha tenido que incluir el argumento “lang” (de language) para que utilice el vocabulario en español (“spanish”). De esta manera, la función busca la presencia de las ocho emociones y los dos sentimientos para cada palabra en nuestro vector.

Una vez obtenidos los sentimientos de las palabras que conforman los *tweets*, se han obtenido diferentes ilustraciones que serán comentadas y adjuntadas en el capítulo 5 de Resultados.

## 5. RESULTADOS.

### 5.1. Resultados de la evolución de la terminología en el NYT.

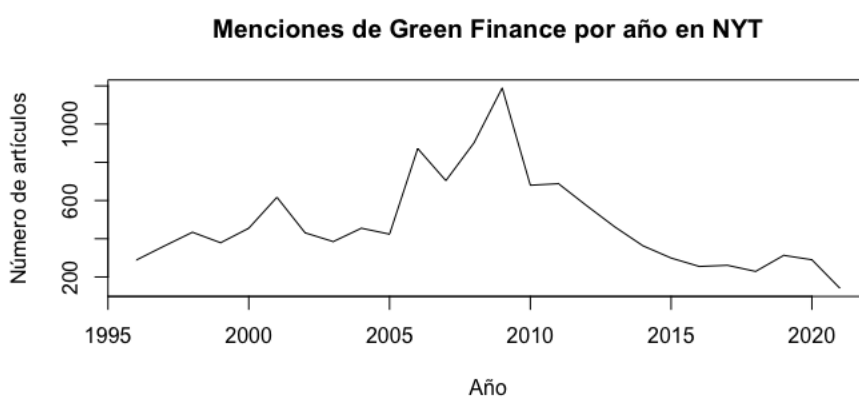
A través del análisis exploratorio descrito en el capítulo 4.2.1 (Menciones en artículos del New York Times) ha sido posible analizar la evolución de las apariciones en artículos del New York Times de distintos términos relativos a las Finanzas Sostenibles.

El primer término analizado ha sido el de “Green Finance”, ya que se ha considerado que puede acumular la máxima información sobre el tema.

Se ha escogido el intervalo temporal que va desde la fecha inicio del 1 de enero de 1996 (un año previo al Protocolo de Kioto de 1997) hasta la fecha final del 1 de abril de 2021.

Mediante la función *nyt\_years\_count*, que aparece en la Ilustración 12, se ha hecho recuento de las apariciones en artículos en el intervalo expuesto, dando un total de 12.425. Tras la contabilización del total de apariciones en artículos del NYT, se han ejecutado diferentes funciones presentadas en el capítulo 4.2.1, y hemos podido ver su evolución a lo largo de los años. En primer lugar, se han obtenido las Menciones de “Green Finance” anuales desde enero de 1996 hasta enero de 2021 (Ilustración 14).

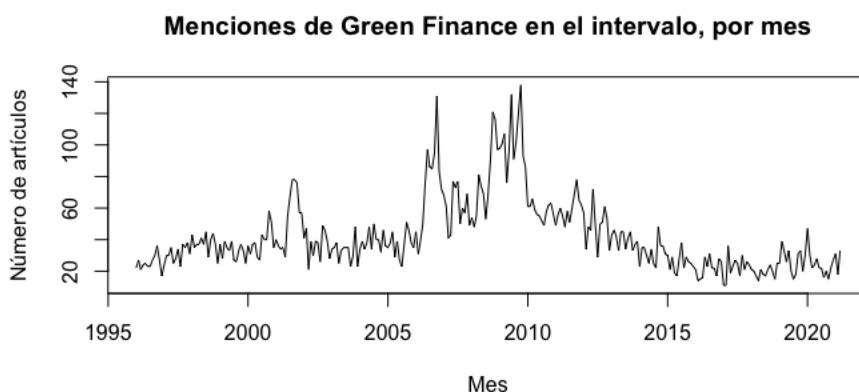
**Ilustración 14: Evolución de las menciones de Green Finance por año en el NYT.**



*Fuente: Elaboración propia en Rstudio*

Para tener una información más completa se ha decidido añadir otro gráfico con la evolución mensual (Ilustración 15).

**Ilustración 15: Evolución de las menciones de Green Finance por mes en el NYT.**



*Fuente: Elaboración propia en Rstudio*

De las ilustraciones anteriores se puede extraer que en el inicio del intervalo en 1995 hay 250 artículos anuales y que experimentan un crecimiento hasta cerca de 500 en el año 1997, lo que pudo deberse al Protocolo de Kioto que aumentó la concienciación general sobre la necesidad de reducir la emisión de gases. La evolución se mantuvo más o menos constante hasta el primer lustro del siglo XXI, cuando se produjo un gran aumento superando los 600 artículos anuales, hecho que puede explicarse por el lanzamiento de los ODM en el año 2000. Sin embargo, sufre una caída que hace que vuelva a la cantidad próxima a 300 artículos anuales, cerca del intervalo entre 200 y 300 que predomina en la mayoría de los años. En el siguiente lustro, de 2005 a 2010 hay mayor volatilidad, siendo generalmente positiva, ya que alcanza el máximo en 2009 llegando a más de 1.000 artículos anuales, lo que puede explicarse, entre otros motivos, por la drástica reducción de inversiones en proyectos de investigación con escasa rentabilidad, como por ejemplo los de la NASA relacionados con exploraciones en el *Deep-space* y en planetas del sistema solar, decantándose las instituciones financieras americanas por proyectos de sostenibilidad en nuestro planeta. Otro motivo probable es el compromiso que adoptó en septiembre de ese año el G20 sobre un "modelo de crecimiento mundial más equilibrado" o el estar sumergidos en una crisis mundial (para muchos la peor desde 1929) que pudo provocar que muchos inversores abandonaran las inversiones en negocios ya maduros con problemas de liquidez y buscaran nuevos proyectos de inversión. Además, en este



mismo año 2009 se creó la CBI, dos años más tarde de la creación de los Bonos Verdes cuya evolución en cuanto a menciones en el NYT será analizada más adelante.

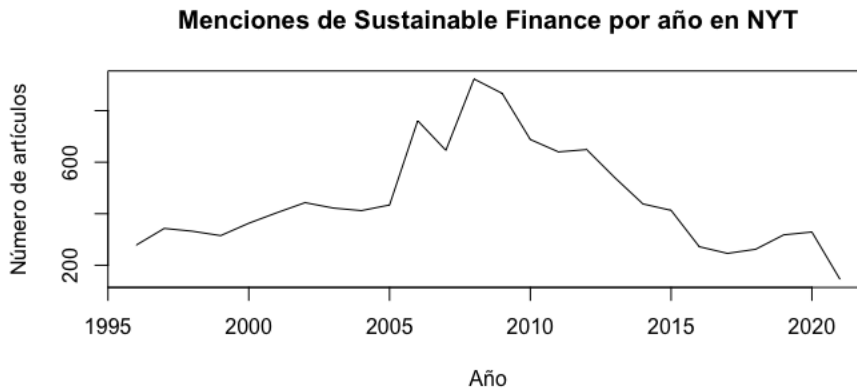
Tras este pico máximo en 2009 caen las menciones significativamente, lo que sorprende especialmente a partir del año 2015 por la aparición sucesiva de diferentes acuerdos y planes adoptados en materia sostenible (Acuerdo de París, la Agenda 2030 y ODS, acuerdos sobre Taxonomía, las recomendaciones del NGFS o el Green Deal). Aun así, conforme nos movemos hacia el año 2020, se percibe un aumento de las menciones, soportado por el hecho del crecimiento que han ido experimentando en los últimos años las finanzas sostenibles, y que, según predicciones, seguirá en los próximos años.

Para tener una idea más completa de la evolución del término de Finanzas Sostenibles, además de analizar la evolución de “Green Finance”, se ha decidido analizar también el término “Sustainable Finance” con la finalidad de observar si el nuevo gráfico recoge una evolución similar en el mismo intervalo temporal entre el 1 de enero de 1996 y el 1 de abril de 2021.

Mediante la función *nyt\_years\_count*, que aparece en la Ilustración 12, se ha hecho recuento que las apariciones en artículos en el intervalo analizado, dando un total de 11.883, cantidad similar, aunque inferior, a la de 12.425 apariciones de “Green Finance”.

Situados en este punto, y para ver si seguían una evolución similar, se obtuvo la Ilustración 16, que muestra la evolución de Menciones de “Sustainable Finance” anuales desde enero de 1996 hasta enero de 2021.

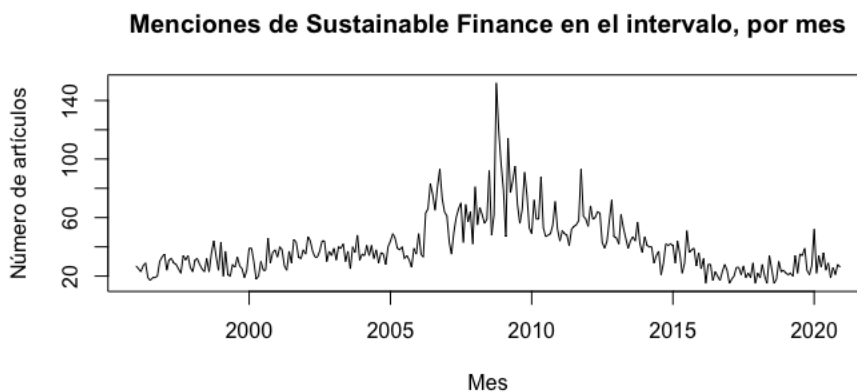
**Ilustración 16: Evolución de las menciones de Sustainable Finance por año en el NYT.**



*Fuente: Elaboración propia en Rstudio*

Al igual que se ha hecho con “Green Finance”, obtenemos la Ilustración 17 sobre las menciones mensuales en el mismo intervalo con el objetivo de tener una información más completa a nuestra disposición.

**Ilustración 17: Evolución de las menciones de Sustainable Finance por mes en el NYT.**



*Fuente: Elaboración propia en Rstudio*

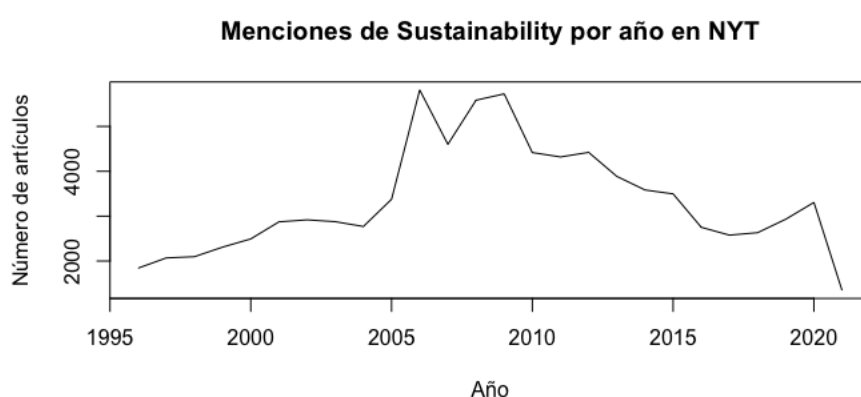
Las ilustraciones sobre la evolución anual y mensual de las apariciones de “Sustainable Finance” en artículos del NYT presentan una distribución prácticamente igual a la de “Green Finance”, algo que cobra sentido al ser dos términos sinónimos.

A su vez, se han construido los gráficos correspondientes a la evolución en el mismo intervalo de tiempo del término “Sustainability”, ya que, al igual que los términos que ya

han sido introducidos, engloba de manera correcta el tema del estudio. Mediante la función *nyt\_years\_count*, que aparece en la Ilustración 12, se ha hecho recuento de las apariciones de este término en artículos en el intervalo expuesto, dando un total de 86.714, un número de apariciones mucho mayor al de los dos términos anteriores.

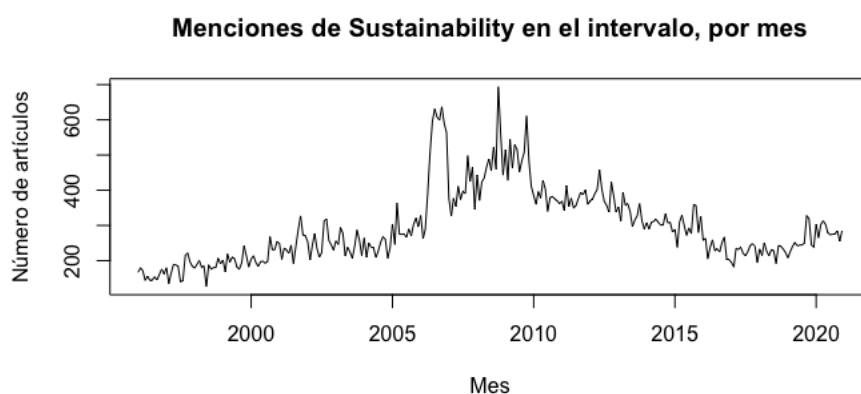
En las ilustraciones 18 y 19 se puede ver la evolución anual y mensual desde el 1 de enero de 1996 hasta el 1 de abril de 2021 del término “Sustainability” en cuanto a sus menciones en el New York Times.

**Ilustración 18: Evolución de las menciones de Sustainability por año en el NYT.**



*Fuente: Elaboración propia en Rstudio*

**Ilustración 19: La evolución de las menciones de Sustainability por mes en el NYT.**



*Fuente: Elaboración propia en Rstudio*

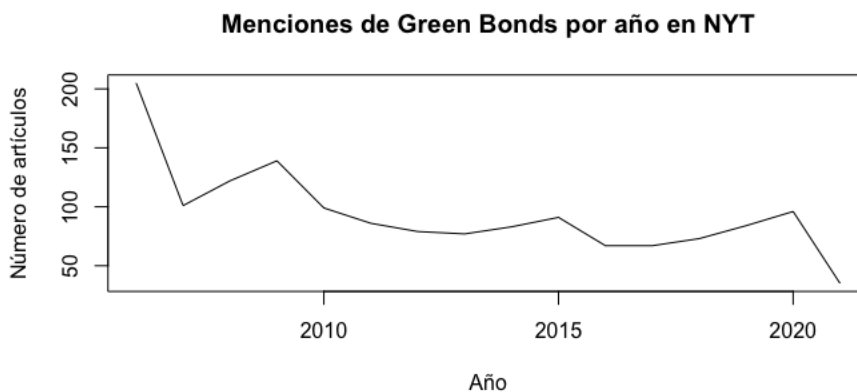
Se puede ver una evolución similar a la de las finanzas sostenibles (“Green Finance” o “Sustainable Finance”), logrando el pico máximo en el segundo lustro del siglo XXI. A partir de ahí se observa una caída hasta el crecimiento que está experimentando en la

actualidad gracias a los acuerdos, a los reglamentos y al impulso que, durante la pandemia que vivimos del COVID-19, se está dando a los proyectos de sostenibilidad.

Otro término analizado ha sido el de “Green Bonds”, haciendo referencia a los bonos verdes, que han sido presentados en el capítulo 2 del trabajo, ya que, como se puede ver en la Ilustración 6, han tenido un volumen de emisiones creciente en la última década. Se quiere analizar si han tenido un impacto similar en los artículos del NYT y si ha seguido una evolución semejante.

Para ello se ha escogido el intervalo temporal diferente que va desde la fecha inicio del 1 de enero de 2006 (un año previo a su creación en 2007) hasta la fecha final del 30 de abril de 2021. Mediante la función *nyt\_years\_count*, que aparece en la Ilustración 12, se ha hecho recuento de las apariciones en artículos en dicho intervalo, arrojando un valor de 1.500 en total. Mas tarde, se ha construido un gráfico para ver las menciones por año de “Green Bonds” en artículos del NYT en el intervalo de tiempo mencionado. El resultado se puede apreciar en la Ilustración 20.

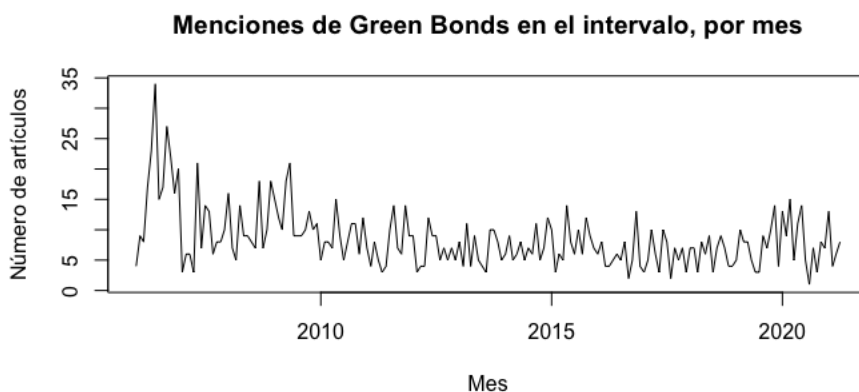
**Ilustración 20: La evolución de las menciones de Green Bonds por año en el NYT.**



*Fuente: Elaboración propia en Rstudio*

Para tener una información más completa se ha decidido añadir la Ilustración 21, que muestra la evolución mensual.

**Ilustración 21: Evolución de las menciones de Green Bonds por mes en el NYT.**



*Fuente: Elaboración propia en Rstudio*

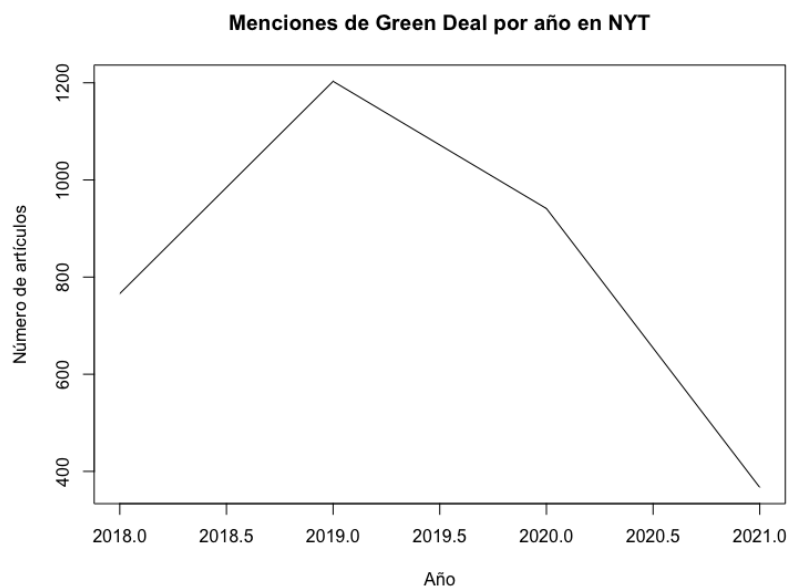
Como se ha expuesto previamente los bonos verdes se crearon en 2007, lo que puede explicar que el mayor volumen de menciones se concentre en el entorno temporal de esa fecha, llegando a más de 30 artículos con menciones directas al mes. A pesar de que el auge de estos instrumentos de financiación se dio en 2013, no se ve un gran crecimiento como el que se dio en su creación, lo que puede llegar a sorprender. Se puede ver que en 2015 aumenta, posiblemente por la aparición de diferentes acuerdos, hasta llegar a su posterior bajada que ha sido revertida en los últimos tiempos, seguramente debido a los mencionados acuerdos y sus acciones, los reglamentos y el impulso durante el COVID-19 en la emisión de Bonos Verdes. Por tanto, se ve que la evolución de las menciones no guarda correspondencia con el volumen de emisiones de Bonos Verdes (ver Ilustración 6).

Se ha decidido incorporar también el término “Green Deal”, que hace referencia al Pacto Verde europeo, plan que pretende lograr la neutralidad climática de la UE para 2050 y avanzar en el cumplimiento de los ODS. El motivo de su inclusión en el trabajo responde a la voluntad de considerar en el análisis el término de un plan específico sobre la materia. Al ser anunciado en 2019, se ha añadido como fecha de inicio enero de 2018 (año previo a su anuncio) y 2021 como fecha final, año que, a pesar de no haber concluido en el momento de la elaboración de este trabajo, permitirá recoger las menciones hasta el 29 de mayo de 2021. Mediante la función *nyt\_years\_count*, que aparece en la Ilustración 12, se ha hecho recuento de las apariciones en artículos en el intervalo expuesto, dando un total de 3.277 hasta el 29 de mayo de 2021, repartiéndose este dato de la siguiente manera:

766 menciones en 2018; 1203 en 2019; 941 en 2020; y 367 en lo que llevamos de año 2021.

A continuación, y siguiendo el proceso adoptado en los demás términos, se ha construido un gráfico para ver las menciones por año de “Green Deal” en artículos del NYT desde enero de 2018 hasta 2021. El resultado se puede apreciar en la Ilustración 22.

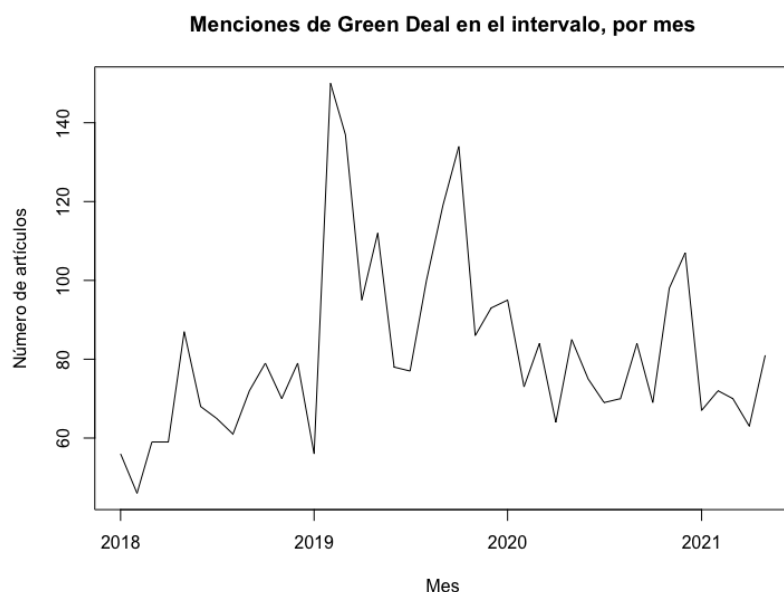
***Ilustración 22: Evolución de las menciones de Green Deal por año en el NYT.***



*Fuente: Elaboración propia en Rstudio*

Al igual que en el resto de los análisis, para tener una información más completa, se ha añadido otro gráfico con la evolución mensual, que se puede ver en la Ilustración 23.

**Ilustración 23: Evolución de las menciones de Green Deal por mes en el NYT.**



*Fuente: Elaboración propia en Rstudio*

Resulta del todo coherente que tuviese su mayor pico de menciones en el año 2019, fecha en la que se produjo su anuncio definitivo. El descenso aparente en 2021 hay que interpretarlo teniendo en cuenta que solo recoge las menciones hasta el mes de junio por lo que se puede considerar que, a final de año, su evolución se mantendrá al alza.

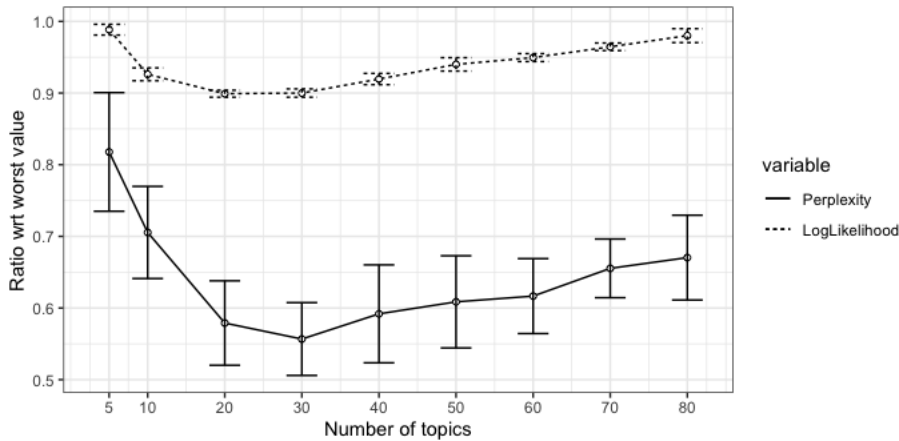
A modo de conclusión sobre la evolución de menciones en artículos del New York Times mostradas en las diferentes ilustraciones obtenidas, se puede decir que generalmente no existe una correlación positiva entre la popularidad e importancia de las finanzas sostenibles en su conjunto y su evolución en los artículos del medio de comunicación escogido, siendo el elemento más significativo para este razonamiento que a pesar de que el año 2021 ha sido de gran popularidad para todos los términos relacionados con la sostenibilidad, no encontramos en dicho año ningún pico de menciones en los términos más globales del tema (“Green Finance”, “Sustainable Finance” y “Sustainability”). En los otros dos términos estudiados (“Green Bonds” y “Green Deal”) hay diferencias en cuanto a su interpretación, ya que en el primero se ha visto que tampoco tiene una correlación positiva con la evolución del volumen de emisiones que existe (ver Ilustración 6), y en el segundo de los términos estudiados, al ser el que ha sido sometido a un intervalo





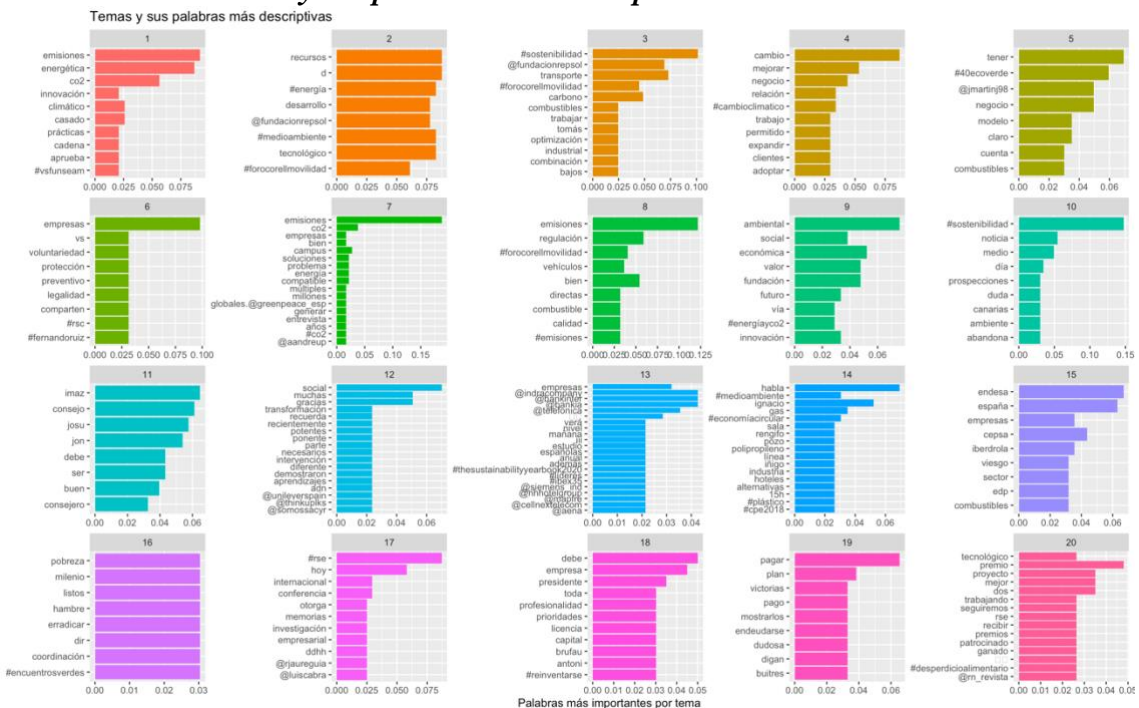
Tras obtener las palabras con mayor frecuencia, se ha construido diferentes gráficos (Ilustración 26) con los temas más relevantes y sus palabras más importantes. Para ello, previamente se calculó el número óptimo de temas (*clusters*) con la “Regla del codo”, que se puede apreciar en la ilustración 25, y que nos proporcionó el valor de 20 como el número óptimo de temas.

**Ilustración 25: Regla para calcular el número óptimo de temas o clusters**



Fuente: Elaboración propia en Rstudio. Datos extraídos de Twitter

**Ilustración 26: Temas y sus palabras más descriptivas**



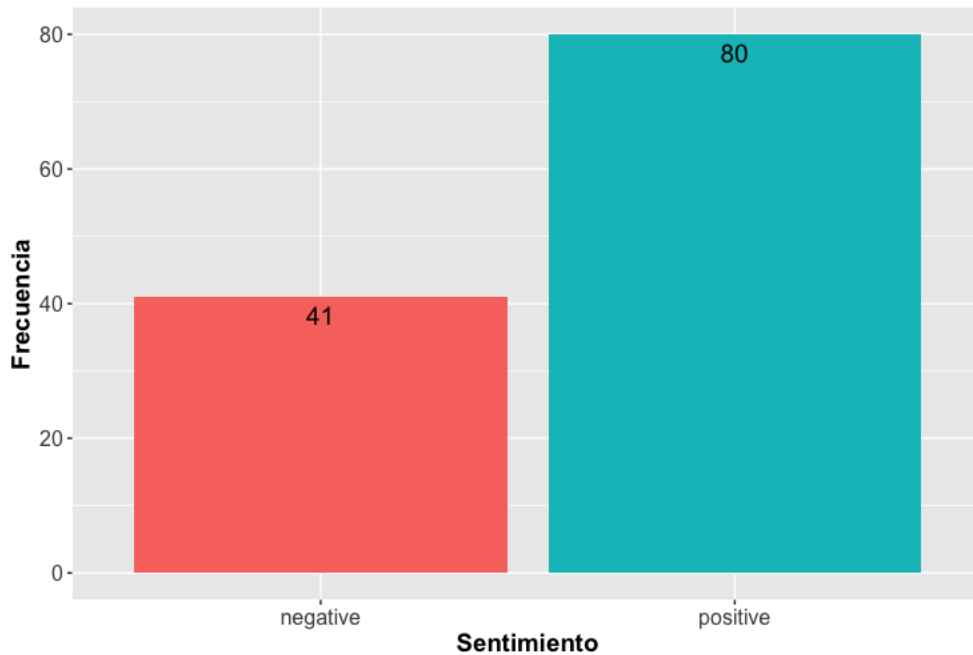
Fuente: Elaboración propia en Rstudio. Datos extraídos de Twitter.

Como ya se comentó en el epígrafe 4.2.2, una vez obtenidos los sentimientos de cada término, se han ejecutado diferentes gráficos que nos aportan información del contenido

de los *tweets* en cuanto a sus sentimientos (positivo y negativo) y emociones (enfado, anticipación, disgusto, miedo, alegría, tristeza, sorpresa y confianza).

En primer lugar, se ha obtenido un gráfico de barras (Ilustración 27) para ver cuál es el sentimiento predominante.

**Ilustración 27: Frecuencia de sentimientos**

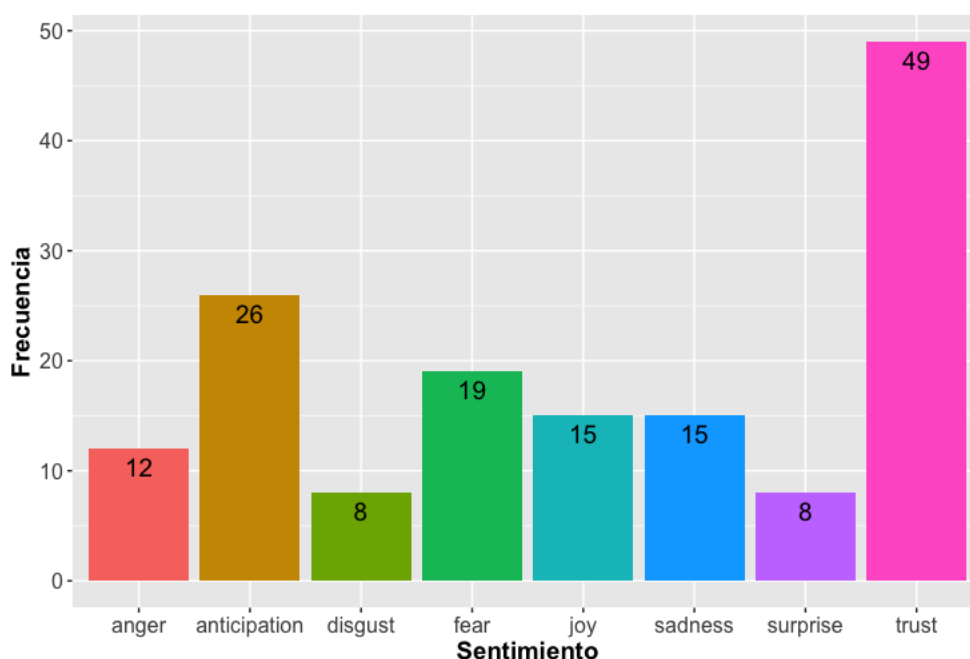


*Fuente: Elaboración propia en Rstudio. Datos extraídos de Twitter.*

Como se puede ver en la previa Ilustración 27 el sentimiento predominante es el positivo, lo que nos da pie a considerar que el público de Twitter inspira mayor positividad que negatividad con relación a la gestión de Repsol sobre la sostenibilidad, información que será desarrollada a continuación. Como se puede observar en la previa Ilustración 27, el sentimiento predominante es el positivo, lo que nos da pie a considerar que el público de Twitter inspira mayor positividad que negatividad con relación a la gestión de Repsol sobre la sostenibilidad, información que será desarrollada a continuación.

En segundo lugar, se ha realizado otro gráfico de barras (Ilustración 28) para ver cuáles son las emociones con mayor presencia en el texto mediante la función *barplot()* y con el procesado de la función *prop.table()*.

**Ilustración 28: Gráfico de barras para ver las emociones con mayor presencia.**



*Fuente: Elaboración propia en Rstudio. Datos extraídos de Twitter.*

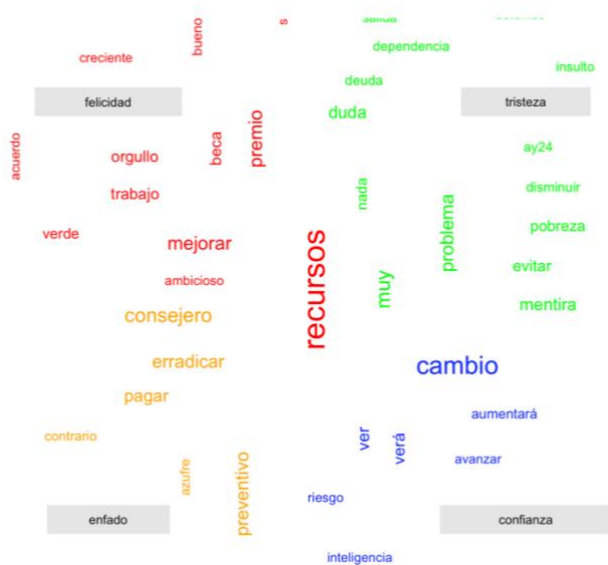
Observando la Ilustración 28, se puede ver que el sentimiento que prevalece es el de confianza (*trust*), seguido del de anticipación (*anticipation*), de miedo (*fear*) y de alegría (*joy*) y tristeza (*sadness*), siendo los demás sentimientos menos destacados en el texto. Sorprende que la confianza sea el que destaca, lo que puede interpretarse como que el público de Twitter que ha mencionado “Repsol” y “Sostenibilidad” tiene confianza en que la compañía española logre avanzar hacia una actividad más sostenible en el futuro al ser la emoción más repetida, y que las diferentes regulaciones van a hacer que Repsol se adapte. Del análisis de la Ilustración 28 se desprende también que otras emociones más “negativas” cuentan con menor peso.

Una vez se ha conocido que la emoción más repetida es la confianza, se procederá a conocer qué palabras son usadas con mayor frecuencia y estén identificadas con la emoción de “confianza”. Dejando de lado palabras que se han considerado como *stopwords* (por ejemplo: director y consejero), los términos más repetidos asociados a la confianza han sido: “mejorar”, “valor”, “premio” y “responsable”. A su vez, el número de palabras únicas con esta emoción alcanza las 49.

Tras obtener los gráficos de barras previos, se ha considerado una buena opción la de construir una nube de emociones con el objetivo de ver las palabras que corresponden a cada emoción. Para ello, se crea primero en *RStudio* un vector en el que se guardan todas las palabras que, en las columnas tras el símbolo \$ tienen un valor mayor a 0.

Se genera un nuevo objeto de tipo vector, que contiene un elemento para el listado de cada emoción. Una vez obtenido el vector, se ha creado un corpus de palabras que asigna un nombre a cada uno de los grupos de palabras, que son: tristeza, felicidad, enfado y confianza. Posteriormente, y aplicando las funciones *set.seed()* y *comparison.cloud* (del paquete *wordcloud*, que ha sido descargado) se ha obtenido la Ilustración 29.

**Ilustración 29: Nubes de emociones**



*Fuente: Elaboración propia en Rstudio. Datos extraídos de Twitter.*

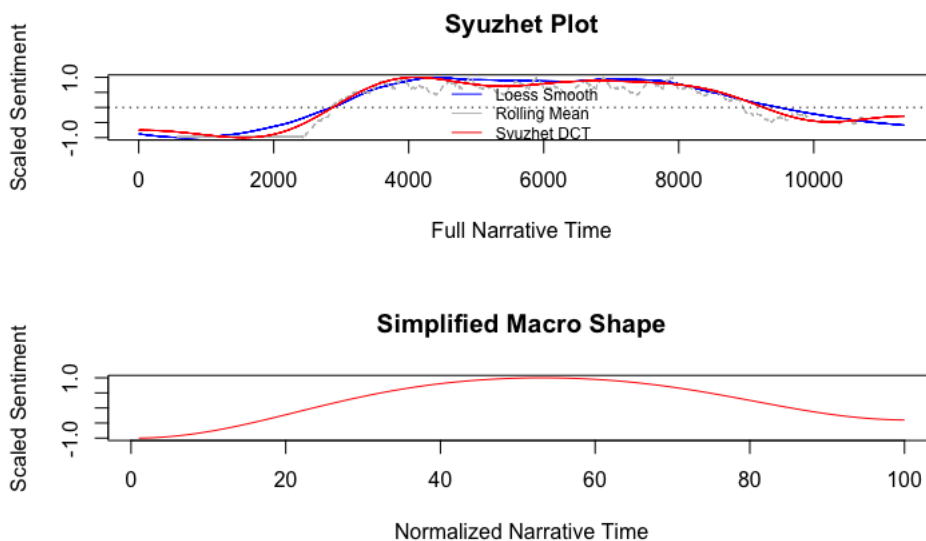
Esta nube de palabras nos muestra términos asociados a las cuatro emociones que hemos especificado. Sorprende alguno de ellos, como los de “ay24”, “muy” o “consejero” que serán fruto de alguna de las consideraciones que se puntualizaron en el epígrafe 4.2.2. como posibles limitaciones del procedimiento.

Esta nube aporta un gran valor visual pero no de forma cuantificada, lo que limita su interpretación, aunque sí nos permite entender que tipo de términos se asocia a cada emoción.

Una vez se ha obtenido una lectura aislada de las emociones mediante la nube, se ha decidido obtener unas gráficas que muestren la fluctuación o evolución de dichas emociones a lo largo del intervalo temporal del estudio sobre los *tweets* relativos a Repsol. Ha sido posible asignando un valor +1 para un momento más positivo y -1 a los más negativos, de manera que calculamos la valencia de los tweets multiplicando los valores de la columna de los valores negativos de nuestro *data frame* por -1 y, tras ello, sumar el valor de la columna de los positivos.

El resultado se puede apreciar en la Ilustración 30.

**Ilustración 30: Evolución de sentimientos en los tweets, siendo Syuzhet Plot la gráfica con todas las medidas del algoritmo y Simplified Macro Shape el gráfico normalizado del anterior.**



*Fuente: Elaboración propia en Rstudio. Datos extraídos de Twitter.*

Aquí, lo que se muestra es que comienzan de forma muy negativa, hasta que toma valores centrales en la primera quinta parte del gráfico, momento desde el que comienza a evolucionar positivamente ( $>0$ ) hasta finalizar el intervalo de forma negativa, pero más neutra que al inicio, tomando valores cercanos al -1.

### 5.3. Análisis del precio de cotización de Repsol

*Ilustración 31: Precio de cierre ajustado de las acciones de Repsol entre 2015 y 2020.*



*Fuente: Elaboración propia en Rstudio. Datos extruidos de Yahoo Finance.*

En la Ilustración 31 se muestra la evolución del valor de cierre de la cotización de Repsol en el periodo 2015 – 2020. Un análisis somero deja ver dos caídas muy fuertes de la cotización, que marcan los periodos de inicio de recuperación acontecidos a sus respectivos términos.

La primera de ellas se produjo a lo largo del año 2015 y tuvo su origen en la caída generalizada de precios de las materias primas como consecuencia del descenso acusado de la demanda en China, país consumidor del 35% de las mismas y por lo tanto con gran capacidad de influenciar las cotizaciones internacionales. En el caso particular del barril de petróleo, a este fenómeno de caída de demanda en China se superpone el encarecimiento sufrido en ese mismo periodo del dólar. La sensibilidad de las compañías petroleras a la cotización del barril de petróleo es un hecho sistemático y esto produjo la caída de la cotización bursátil de Repsol durante buena parte del 2015. Por otro lado, la aparición del fraude “Dieselgate” consistente en la falsificación de las emisiones oficiales de motores diésel del grupo Volkswagen en el 2015, ha lastrado en general desde el 2015 las cotizaciones de todas las petroleras en tanto en cuanto las ha expuesto de manera radical a la transición energética asociada a la movilidad.

La segunda caída de produjo en el año 2020 y en este caso el factor que lo explica de forma mayoritaria es la contracción de la demanda de energía y movilidad a partir del mes de marzo como consecuencia de la aparición global del virus COVID-19.

Estos hechos atenúan el posible efecto que hayan tenido los *tweets* sobre la cotización de la empresa Repsol, y deben tenerse en cuenta a la hora de interpretar el posible impacto que han tenido los sentimientos extraídos de los *tweets* sobre la cotización de la petrolera, que será analizado en el epígrafe 5.4.

Previo al análisis del impacto de la opinión del público de Twitter sobre la cotización de Repsol, resulta apropiado analizar la exposición de la empresa en términos de riesgo.

Se han tenido en cuenta los índices de riesgo ESG de Sustainalytics, que evalúan el grado en que el valor del negocio empresarial de una empresa está en riesgo a causa de problemas de medioambiente, responsabilidad social y gobierno corporativo. Sustainalytics es una organización que aporta una base de datos PAI<sup>38</sup> para apoyar a los inversores en el cumplimiento del SFDR<sup>39</sup> de la UE.

El Índice de riesgo ESG total es de 23, lo que supone un valor medio de riesgo en comparación con otras empresas. Una vez se descompone ese valor entre sus tres diferentes componentes, se descubre que obtiene una nota de 10 en cuanto al riesgo medioambiental, que es el que tiene una mayor importancia en el estudio. Por tanto, en términos medioambientales, se puede destacar a Repsol como una de las más sensibles. Para completar los datos, obtiene un 6,8 en la puntuación de riesgo social y un 6 en puntuación de riesgo de gobernanza, que suponen unos valores más bajos que hacen que el riesgo total ESG de Repsol se mantenga en un nivel medio. Estos *ratings* se pueden ver en la Ilustración 32.

---

<sup>38</sup>PAI son las siglas de *Principal Adverse Impact*, que hacen referencia a los efectos potencialmente negativos sobre los factores de sostenibilidad que están relacionados con las decisiones de inversión.

<sup>39</sup>SFDR son las siglas de *Sustainable Finance Disclosure Regulation*, que es el Reglamento sobre Divulgación de Finanzas Sostenibles que exige a las gestoras de activos incluir los PAI en sus decisiones de inversión.

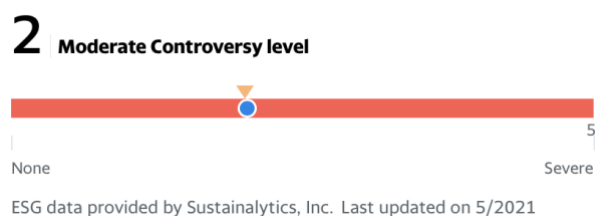
**Ilustración 32: Extracto de los ratings de Sustainalytics sobre Repsol en relación con su calificación sobre los riesgos ESG**



Fuente: pantallazo de la web de Yahoo Finance, que extrajo los datos de Sustainalytics.

También se ha considerado integrar el nivel de controversia que le aplica Sustainalytics, que se refiere al potencial impacto negativo que una empresa implicada en incidentes puede provocar a sus *stakeholders*, al medio ambiente y a la misma compañía internamente. Estas controversias se puntúan sobre una escala del uno al cinco, en orden creciente en cuanto a gravedad e impacto potencial. Repsol obtiene una puntuación de 2 (puntuación obtenida en el mes de mayo de 2021), lo que la sitúa en el promedio de empresas similares. Se puede ver esta calificación en la Ilustración 33.

**Ilustración 33: Nivel de controversia de Repsol según los criterios de Sustainalytics**



Fuente: pantallazo de la web de Yahoo Finance, que extrajo los datos de Sustainalytics.

Además, Repsol presenta un índice referente a la exposición a materia ESG con un resultado *high*, el más alto. Sin embargo, se comprueba que la dirección de la compañía está gestionando de una manera *strong* todos los temas ESG. Se pueden observar ambas calificaciones en la Ilustración 34.



### ***Ilustración 34: Extracto de los ratings de Sustainalytics sobre Repsol en materia de exposición al riesgo y gestión del riesgo***



*Fuente: Sustainalytics*

Por tanto, se puede concluir que Repsol es una compañía que, a pesar de estar gestionada de una manera correcta para combatir los riesgos, es muy sensible ante los impactos medioambientales, teniendo una alta exposición que puede provocar gran volatilidad en el precio de sus acciones frente a diferentes hechos ambientales o de criterios ESG.

Se verá a continuación si, además de ser influida por hechos ambientales, es sensible también a los sentimientos que emiten los *tweets* sobre la compañía en relación con la sostenibilidad.

#### **5.4. Análisis de correlación entre sentimiento promedio de los tweets y el precio de cotización de Repsol**

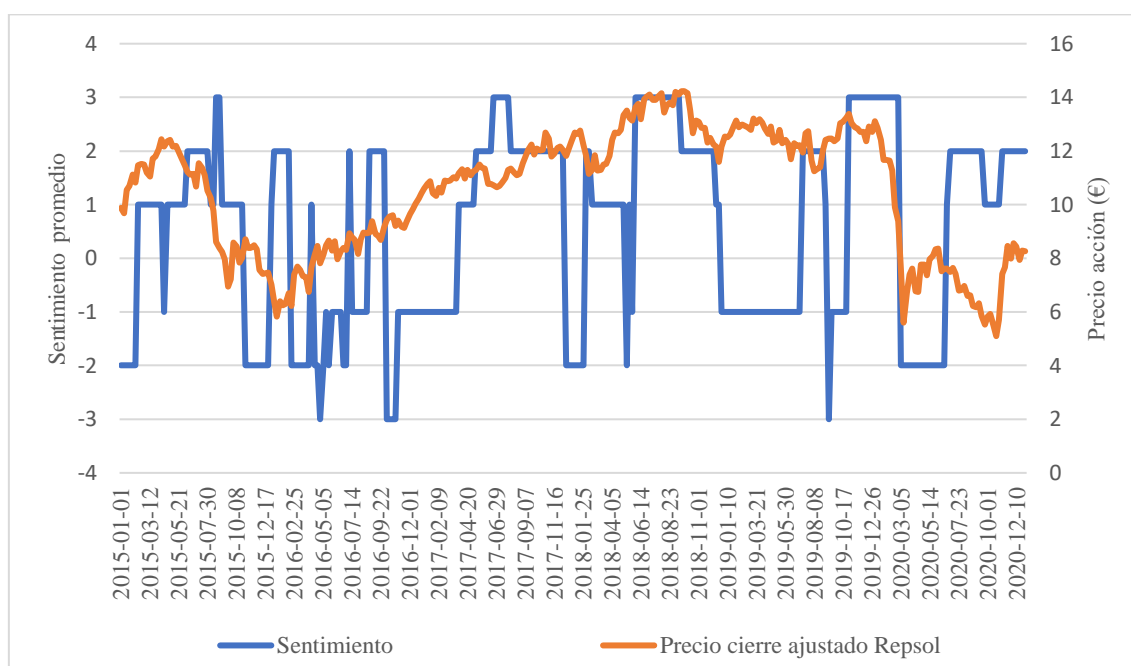
Uno de los objetivos de este trabajo era analizar si existía correlación entre el sentimiento extraído de los *tweets* que hablan de Repsol con referencia a la sostenibilidad (como ya se ha dicho previamente, es una empresa que tiene una alta exposición a ello), y la cotización de la empresa petrolífera en bolsa. Por ello, se ha considerado oportuno elaborar la Ilustración 35 para ver si existe dicha correlación o si, por el contrario, los *tweets* no han tenido impacto en los movimientos bursátiles de la empresa española.

Para ello, en primer lugar, se han obtenido los sentimientos promedios semanales de los *tweets* extraídos para que tengan la misma estructura temporal que las cotizaciones descargadas, que recordamos también eran semanales. Se han obtenido estos sentimientos promedios usando el *lexicon Afinn*, que clasifica los sentimientos en una escala que va desde el -4, denotando el sentimiento más negativo, hasta el +4, denotando el sentimiento más positivo, en función de las distintas ocho emociones ya descritas en el epígrafe 4.2.2.

Una vez se ha conseguido lo anterior, se exportaron dichos sentimientos promedios semanales a Excel para la obtención del gráfico.

En segundo lugar, las cotizaciones semanales de Repsol han sido también exportadas a Excel para así dibujar el gráfico de dos ejes que vemos en la Ilustración 35.

**Ilustración 35: Correlación entre el sentimiento promedio semanal de los tweets y el precio de cierre ajustado semanal en el intervalo entre 2015 y 2020.**



*Fuente: Elaboración propia en Excel. Datos extraídos de Yahoo Finance y Twitter.*

Observando la Ilustración 35, se puede concluir que no existe ningún tipo de correlación entre el valor del precio ajustado de las acciones bursátiles de Repsol y el sentimiento promedio obtenido de los *tweets* que mencionan “Repsol” y “Sostenibilidad” de forma concurrente. Se determina que no hay correlación puesto que hay tramos, como el que va de diciembre de 2016 hasta mayo de 2017, o también entre junio y octubre de 2018 (en el que se dio la cotización más alta del periodo analizado) donde hay una visible correlación positiva y otros, como puede ser el de julio a octubre de 2020, en el que se da una correlación inversa o negativa.

En los casos que hay una correlación negativa, especialmente cuando el sentimiento promedio es positivo y hay movimientos bajistas en su cotización, se pueden deber al tipo de mensaje que se emite en Twitter, ya que en ocasiones son emitidos por miembros de la misma empresa que suelen realizar comunicados positivos en situaciones adversas para la compañía, pudiendo este hecho explicar ese estado más optimista. También resulta interesante recordar que, en el previo análisis de sentimiento, las emociones positivas superaban con creces a las negativas.

Se podría por lo tanto concluir que los mensajes emitidos en Twitter acerca de la empresa Repsol con relación a la sostenibilidad no han influido directamente en su cotización bursátil, sino que habrá estado determinada por otros factores que hayan causado estas oscilaciones, visibles en la Ilustración 31 y expuestos en el epígrafe 5.3. que trata el análisis del precio de cotización de Repsol.

## 6. CONCLUSIONES

El presente trabajo de investigación ha tratado de evidenciar la relevancia que está teniendo el tema de análisis, que son las finanzas sostenibles, en los últimos años mediante el uso de alguna de las técnicas del *Big Data*.

El *Data Analytics* es una herramienta muy potente que resulta de aplicación a un número muy elevado de análisis, y que puede facilitar la toma de decisiones empresariales, desde las relativas a lanzamiento de productos y perspectivas de flujos de caja, hasta la medición de repercusiones en distintas métricas de la actividad empresarial. En nuestro caso la hemos aplicado para ver impactos en la cotización bursátil de una empresa, pero el campo de aplicación es casi infinito.

El presente estudio ha presentado además diferente información sobre la evolución del tema de análisis, contrastándola con la experimentada por las menciones en un prestigioso medio de comunicación como el New York Times, estudiando además el impacto que ha tenido su auge en los últimos años en la cotización bursátil de una empresa que está altamente expuesta a los riesgos medioambientales y climáticos como es Repsol.

La sostenibilidad ha sido una cuestión que ha traspasado fronteras y ha cobrado gran repercusión en el público en general, especialmente a través de la creación de una serie de acuerdos, instrumentos financieros, organismos y leyes que han determinado su inclusión en las agendas de todos los gobiernos y organizaciones empresariales a nivel mundial, y que han sido expuestos en este trabajo para una mejor comprensión de la evolución experimentada por el tema de análisis.

La globalidad alcanzada por las finanzas sostenibles ha motivado el análisis de la evolución de diferentes términos relacionados en el New York Times, medio escogido por su relevancia mundial, por la posibilidad de utilizar su base de datos y por ser un referente en la información y en la opinión.

Tras el análisis realizado sobre la evolución de la terminología sostenible, medida en menciones en artículos, se ha comprobado que a pesar del gran volumen de información encontrado sobre ellos y que presumiblemente tienda al alza en el futuro, los resultados

obtenidos no reflejan al completo el enorme crecimiento que están experimentando las finanzas sostenibles desde el año 2007 en que se crearon los bonos verdes, y especialmente a lo largo de la última década.

Se ha observado en los análisis de los tres términos que engloban mejor el asunto de la sostenibilidad (“Green Finance”, “Sustainable Finance” y “Sustainability”), que distintos hechos cristalizaron en el aumento del número de menciones hasta lograr su valor máximo a finales de la primera década del siglo XXI. A partir de ese momento se produjo un descenso, tendencia que se ha revertido en los últimos dos años en los que se presenta una línea ascendente como consecuencia de la aparición del Green Deal, la NGFS, el Plan de Acción para vincular el sistema financiero al desarrollo sostenible o el NextGenerationEU, y que puede augurar grandes aumentos de número de menciones en los próximos años. En cuanto al término “Green Bonds” se ha visto que experimentó su mayor número de menciones después de su creación en el 2007, tuvo un descenso a partir de ese momento que ha revertido en los últimos dos años. Por último, en relación con el “Green Deal”, se ha observado un valor máximo de menciones tras su creación en el 2019, y que, aunque esté en descenso, el número de artículos que tratan el término se está moviendo entre 50 y 150 artículos al mes.

Está claro que la “revolución sostenible” ha tenido una gran influencia en las cotizaciones de diversidad de empresas en todo el mundo, y es por ello por lo que se ha tratado de ver el impacto en Repsol, utilizando herramientas de *data analytics* y extrayendo sentimientos de los datos de Twitter, una red social que resulta de gran relevancia por la posibilidad que brinda para expresar noticias y opiniones.

Una vez realizado el estudio cualitativo y analítico, se ha podido concluir que los sentimientos de Twitter respecto al tema que se ha analizado no presentan ningún tipo de correlación con el precio bursátil de Repsol. Esta conclusión, sin embargo, no significa que el precio de las acciones de Repsol no haya sido afectado por la expansión de las finanzas sostenibles, ya que los sentimientos de Twitter que se han tratado han sido fruto de una muestra pequeña al ser un tema que no acostumbra a ser dialogado en la red social Twitter por lo que se ha podido comprobar. Aun así, se han podido sacar diferentes conclusiones sobre los sentimientos que suscita la gestión sostenible de la compañía española. Se pudo extraer que los términos más repetidos entre los *tweets* que conforman

la muestra son referentes a las emisiones, a la competencia, al factor tecnológico o al propio negocio. Además, el sentimiento predominante en los *tweets* fue el positivo, mostrando especial relevancia la emoción de confianza, lo que ha dado lugar a que se concluya que el público de Twitter tiene una reacción positiva que confía en las acciones que está realizando Repsol. Una interpretación consecuente con la labor de la empresa, preocupada por tener una actividad más sostenible. También podemos apuntar como “drivers” que han influido en los resultados de las métricas adoptadas, el hecho de que el incremento de ventas que sin ambages resulta como efecto directo de la imagen de marca de Repsol y de su preocupación por la sostenibilidad del medio ambiente, aunque repercuta inequívocamente en la cuenta de resultados de la compañía no se ve acompañada de variaciones notables en la cotización de Repsol.

El trabajo de investigación pretende sentar las bases para otros análisis similares en el futuro apasionante que las finanzas sostenibles nos aguarda. Aun así, ha estado sujeto a una serie de limitaciones que serán presentadas a continuación.

### **6.1. Limitaciones del estudio**

El principal problema del estudio ha sido la falta de datos, especialmente para el análisis de sentimiento realizado para el que se ha contado con una muestra de 344 *tweets* que, como ya se dijo, resulta muy baja en comparación con los millones de *tweets* que se emiten al día. Este hecho ha sido el resultado no solo del bajo volumen de diálogo sobre el tema (Repsol y sostenibilidad en conjunto) sino también de la posible eliminación de *tweets* por parte de Twitter o de los usuarios.

Otro inconveniente con el que hemos tenido que lidiar ha sido la imposibilidad de crear grafos para estudiar la relevancia o influencia de cada uno de los “nodos” (usuarios) sobre la cotización, ya que no es comparable el tener 344 *tweets*, que tener 344 *tweets* y sus respectivas influencias, al ser igualmente apreciable la calidad de los *tweets* como su volumen.

Otra limitación principal ha estado relacionada directamente con la utilización del análisis de sentimientos, herramienta que, por naturaleza, no es exacta y está sujeta a potenciales restricciones, siendo alguna de ellas: la problemática de interpretar recursos retóricos

(aunque en nuestro caso específico no haya supuesto un gran desafío); el hecho de que el léxico en español sea una traducción directa realizada por traducción automática puede entorpecer una interpretación adecuada (aunque la traducción al español sea una de las más fiables); o el procesamiento de la negación.

Respecto a la cotización bursátil de las acciones de Repsol y su evolución, se ha encontrado como limitación el hecho de que sus dos mayores caídas en bolsa hayan estado motivadas por dos hechos ajenos a la sostenibilidad, y que por su contundencia e importancia han difuminado el impacto que las opiniones de Twitter han podido tener, y también han dificultado la extracción de conclusiones sobre la correlación de la cotización con dichos sentimientos de Twitter. Además, al haber escogido los valores en el Mercado Continuo Español, se ha podido obviar la consecución de diferentes resultados si se hubiese escogido la cotización en USA, aunque se considera que los arbitrajes entre las diferentes cotizaciones no son lo suficientemente potentes como para alterar los resultados de este estudio.

## 7. BIBLIOGRAFÍA

- Alto Comisionado para la Agenda 2030. (2018). *Alto Comisionado para la Agenda 2030*.  
[https://transparencia.gob.es/transparencia/transparencia\\_Home/index/MasInformacion/Informes-de-interes/Sociedad\\_y\\_bienestar/Alto-Comisionado-Agenda-2030.html](https://transparencia.gob.es/transparencia/transparencia_Home/index/MasInformacion/Informes-de-interes/Sociedad_y_bienestar/Alto-Comisionado-Agenda-2030.html)
- Aslam, S. (2021, 6 enero). • *Twitter by the Numbers (2021): Stats, Demographics & Fun Facts*. Omnicore. <https://www.omnicoreagency.com/twitter-statistics/>
- Banco de España. (2019, enero). *INNOVACIÓN FINANCIERA PARA UNA ECONOMÍA SOSTENIBLE* (Documentos Ocasionales N.o 1916).  
<https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSerias/DocumentosOcasionales/19/Fich/do1916.pdf>
- Brandwatch. (2015, 10 febrero). *Entendiendo el análisis de sentimiento: qué es y para qué se usa*. Obtenido el 1 de junio de 2021 de Brandwatch:  
<https://www.brandwatch.com/es/blog/analisis-de-sentimiento/>
- Cabrera, M. (2018, 14 marzo). *La importancia de Twitter y los mercados*. BBVA NOTICIAS. <https://www.bbva.com/es/la-importancia-de-twitter-y-los-mercados/>
- Carbon Tracker. (2020, 3 Junio). *Mind The Gap: the \$1.6 trillion energy transition risk*. Carbon Tracker Initiative. <https://carbontracker.org/reports/mind-the-gap/>
- Climate Bonds Initiative. (2019, octubre). *Taxonomía de Climate Bonds Initiative UNA GUÍA DE ACTIVOS Y PROYECTOS ALINEADOS AL CLIMA*.  
<https://www.climatebonds.net/files/files/CBI-Taxonomy-Full-Spanish-Oct19%20Final.pdf>
- CNMV. (s. f.). *CNMV - Plan de Acción de la Comisión Europea*. Recuperado 3 de junio de 2021, de <https://www.cnmv.es/portal/Finanzas-Sostenibles/Plan-accion-CE.aspx>
- CNMV. (s. f.). *Qué son*. <https://www.cnmv.es/portal/Finanzas-Sostenibles/Que-son.aspx?lang=gl>. Recuperado 6 de mayo de 2021, de <https://www.cnmv.es/portal/Finanzas-Sostenibles/Que-son.aspx?lang=gl>



- Comisión Europea (2019, 12 octubre). *Un Pacto Verde Europeo*.  
[https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal\\_es](https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal_es)
- Comisión Europea. (2020a, 11 noviembre). *Plan de recuperación para Europa*.  
[https://ec.europa.eu/info/strategy/recovery-plan-europe\\_es](https://ec.europa.eu/info/strategy/recovery-plan-europe_es)
- Comisión Europea. (2020b, 6 marzo). *Mecanismo para una Transición Justa: garantizar que nadie se quede*. [https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal/actions-being-taken-eu/just-transition-mechanism\\_es](https://ec.europa.eu/info/strategy/priorities-2019-2024/european-green-deal/actions-being-taken-eu/just-transition-mechanism_es)
- Comisión Europea. (2020c, marzo 7). *Overview of sustainable finance*.  
[https://ec.europa.eu/info/business-economy-euro/banking-and-finance/sustainable-finance/overview-sustainable-finance\\_en](https://ec.europa.eu/info/business-economy-euro/banking-and-finance/sustainable-finance/overview-sustainable-finance_en)
- Comisión Europea. (s. f.). *Members and Observers of the Platform on Sustainable Finance Overview of plenary and subgroups*.  
[https://ec.europa.eu/info/sites/default/files/business\\_economy\\_euro/banking\\_and\\_finance/documents/eu-platform-on-sustainable-finance-members\\_en.pdf](https://ec.europa.eu/info/sites/default/files/business_economy_euro/banking_and_finance/documents/eu-platform-on-sustainable-finance-members_en.pdf)
- Consejo Académico de Finanzas Sostenibles. (2020, enero). *LAS FINANZAS SOSTENIBLES. Estado de la cuestión y motivaciones para su desarrollo* (Monografía N°1).  
[https://afiescuola.storage.googleapis.com/cms\\_multimedia/cms\\_medias/files/000/001/079/original/Las-Finanzas-Sostenibles.pdf?1579683160](https://afiescuola.storage.googleapis.com/cms_multimedia/cms_medias/files/000/001/079/original/Las-Finanzas-Sostenibles.pdf?1579683160)
- Deloitte. (2020). *2020 Global Impact Report*.  
<https://www2.deloitte.com/global/en/pages/about-deloitte/articles/global-report-home.html>
- Eckert, E. (2021). *Sustainability in the European Union: Analyzing the Discourse of the European Green Deal*. MDPI.
- Fariña, P. M. (2019, 8 julio). *El papel de los emisores e inversores del sector público en las finanzas sostenibles*. BBVA NOTICIAS.  
<https://www.bbva.com/es/opinion/el-papel-de-los-emisores-e-inversores-del-sector-publico-en-las-finanzas-sostenibles/>

- Fernández, R. (2021, 30 marzo). *Twitter en España - Datos estadísticos*. Statista.  
<https://es.statista.com/temas/3595/twitter-en-espana/>
- Gámez, M. J. (s. f.). *Objetivos y metas de desarrollo sostenible*. Desarrollo Sostenible.  
Recuperado 26 de mayo de 2021, de  
<https://www.un.org/sustainabledevelopment/es/objetivos-de-desarrollo-sostenible/>
- Germanwatch. (2021). *Índice de Riesgo Climático Global 2021*.  
<https://germanwatch.org/sites/default/files/Resumen%20Indice%20de%20Riesgo%20Climático%20Global%202021.pdf>
- González, A. (2019, 18 octubre). *La Comisión Europea lanza una plataforma internacional para coordinar un marco global de finanzas sostenibles*. Spainsif.  
<https://www.spainsif.es/plataforma-internacional/>
- Grupolambda. (2021, 14 abril). *Web Scraping y Text Mining para Economía y Negocios*.  
[https://www.grupolambda.com.pe/featured\\_item/web-scraping-y-text-mining-para-economia-y-negocios/](https://www.grupolambda.com.pe/featured_item/web-scraping-y-text-mining-para-economia-y-negocios/)
- Hearst, M. (1999). *Untangling text data mining*.
- ICMA. (2018). *Los Principios de los Bonos Verdes 2018*.
- Isasi, J. (2021, 23 marzo). *Análisis de sentimientos en R con «syuzhet»*. Programming Historian.  
<https://programminghistorian.org/es/lecciones/analisis-de-sentimientos-r>
- Jáurregui, C. (2018, 19 noviembre). *Análisis de sentimiento en R*. LinkedIn.  
<https://www.linkedin.com/pulse/an%C3%A1lisis-de-sentimiento-en-r-carlos-j%C3%A1urregui-fern%C3%A1ndez>
- MDPI. (2020). *What future of Green Bond Market*.
- Nasukawa. (2001). *Text analysis and knowledge mining system*.
- NGFS. (2019, abril). *Un llamado a la acción El cambio climático como fuente de riesgo financiero* (Primer informe integral).

[https://www.ngfs.net/sites/default/files/medias/documents/reporte-de-la-ngfs-una-llamada-a-la-accion\\_es\\_290719\\_0.pdf](https://www.ngfs.net/sites/default/files/medias/documents/reporte-de-la-ngfs-una-llamada-a-la-accion_es_290719_0.pdf)

OMC. (s. f.). *Los Objetivos de Desarrollo del Milenio de las Naciones Unidas*. Recuperado 14 de mayo de 2021, de [https://www.wto.org/spanish/thewto\\_s/coher\\_s/mdg\\_s/mdgs\\_s.htm](https://www.wto.org/spanish/thewto_s/coher_s/mdg_s/mdgs_s.htm)

ONG Manos Unidas. (2020, 26 junio). *Protocolo de Kioto*. <https://www.manosunidas.org/observatorio/cambio-climatico/protocolo-kioto>

Rayón, A. (2020a). *API y Webscraping: técnicas de enriquecimiento de bases de datos*.

Rayón, A. (2020b). *Text Analytics*. Materiales docentes Diploma en Fundamentos en Business Analytics, Universidad Pontificia Comillas, Madrid.

Repsol (2021, 15 marzo). *Accionistas e inversores de Repsol en el mundo - repsol.energy*.

<https://www.repsol.com/es/accionistas-inversores/informacion-privilegiada-y-relevante/comunicaciones-cnmv/index.cshtml?year=2016&month=all&type=HR;IR;IP>

Repsol. (2020, 11 agosto). *Cotización de la acción de Repsol - repsol.energy*.

<https://www.repsol.com/es/accionistas-inversores/informacion-bursatil/cotizacion-de-la-accion/index.cshtml>

Repsol. (2020). *Hoja de ruta de Repsol ante el cambio climático*.

[https://www.repsol.com/imagenes/global/es/hoja-ruta-repsol-ante-cambio-climatico\\_tcm13-155816.pdf](https://www.repsol.com/imagenes/global/es/hoja-ruta-repsol-ante-cambio-climatico_tcm13-155816.pdf)

Rzezak, J. (2020, 22 de junio). *Text mining con R: Análisis de sentimiento en Twitter*.

*RPosts*. <https://rpubs.com/JonathanRzezak/630972>

Sala de Inversión. (2019, 23 diciembre). *¿Cómo afecta al mercado las noticias? El mercado se alimenta de noticias*. Sala de Inversión.

<https://www.saladeinversion.com/articulos/como-afecta-al-mercado-las-noticias/>

Sánchez, M. G. (2017, 13 de marzo). *La importancia de Twitter y los mercados*.

Sustainalytics (s. f.). *Company ESG Risk Ratings – Sustainalytics*. Recuperado 3 de junio de 2021, de <https://www.sustainalytics.com/esg-rating/repsol-s-a/1008201748>

Sustainalytics. (s. f.) *Sustainable Finance Disclosure Regulation Solutions*. Recuperado 6 de junio de 2021, de <https://www.sustainalytics.com/landing-pages/sustainable-finance-disclosure-regulation-solutions>

TCFD. (2017, junio). *Recomendaciones del grupo de trabajo sobre declaraciones financieras relacionadas con el clima* (Reporte Final). <https://assets.bbhub.io/company/sites/60/2020/10/TCFD-2017-Final-Report-Spanish-Translation.pdf>

TCFD. (s. f.). *Supporters*. Recuperado 3 de junio de 2021, de <https://www.fsb-tcfd.org/supporters/>

UNEP-FI. (2012). *PRINCIPIOS DE BANCA RESPONSABLE*. <https://www.unepfi.org/wordpress/wp-content/uploads/2019/10/1.-Gu%C3%ADa-R%C3%A1pida.pdf>

Vicenteño, J. (2017). Oracle Big Data Discovery: Análisis Exploratorio y descubrimiento de información. *Oracle Technology Network*. Obtenido el 23 de mayo 2021 de: <https://www.oracle.com/technetwork/es/articles/database-performance/big-data-discovery-analisis-3437179-esa.html>

World Commission on Environment and Development. (s.f). *Report of the World Commission on Environment and Development: Our Common Future*. <https://sustainabledevelopment.un.org/content/documents/5987our-common-future.pdf>

## 8. ANEXOS

### 8.1. Código utilizado para el análisis de la evolución de la terminología sostenible en cuanto a menciones en artículos del NYT.

El siguiente extracto muestra el código usado para el caso específico del “Green Deal”, por lo que usa las fechas específicas para dicho término. Para el resto de la terminología sostenible, únicamente habría que cambiar los apartados de fecha (el contenido) y de  $q$ .

```
#Eliminamos todos los objetos en memoria
rm(list = ls())

#Cambiar el directorio de trabajo
setwd(dirname(rstudioapi::getActiveDocumentContext()$path))
getwd()

# Cargamos las librerías que vamos a necesitar
library(httr) # Para hacer llamadas a una API
library(jsonlite) # Decodificar el resultado que viene en un JSON a
un Dataframe

# API de New York Times: https://developer.nytimes.com/
# 1. Nos creamos una cuenta donde pone "Get started"
# 2. Nos creamos una APP
# 3. Cogemos como API Article Search
# 4. Copiamos la API KEY, que nos permitirá autenticarnos en todas
las llamadas

# Lo primero que tenemos que hacer
apikey <- 'k9lukrGJJYPoDhuHTluF13jcHBlahRxd'

# Toda llamada a una API, tiene:
# - base_url: la que venga en la ayuda de la API
# - query: los parámetros. Viene como "q" por defecto.
base_url <-
"https://api.nytimes.com/svc/search/v2/articlesearch.json"
```

```

r <- GET(base_url,
          query=list(q="Green Deal", "api-key"=apikey))
r
# Comprobamos el código que nos devuelve la API.

# Lo guardamos a un objeto de R, para entender mejor lo que devuelve
objeto <- content(r, 'parsed')
class(objeto)
names(objeto)
objeto$status
names(objeto$response) # the actual data
objeto$response$meta # metadata

# Cargamos la librería que necesitamos para decodificar la respuesta
library(jsonlite)
# Otra manera de procesar la respuesta: estructurar el JSON
# Y ahora leemos el JSON que nos devuelve
respuesta <- fromJSON(content(r, "text"))
# Podemos preparar un dataframe a partir de la respuesta
datos<-data.frame(respuesta$response$docs$web_url,
                  respuesta$response$docs$snippet,
                  respuesta$response$docs$section_name)

# Vamos a usar algunas facetas para hacer la consulta: las fechas
r <- GET(base_url, query=list(q="Green Deal",
                              fq="section_name: ('Business')",
                              "api-key"=apikey,
                              "begin_date"=20180101,
                              "end_date"=20210101))
r
respuesta <- fromJSON(content(r, "text"))
respuesta$response$meta

json <- content(r, 'parsed')
json
json$status
json$response$meta

```

```

# Nos hacemos una función para contar el número de apariciones
nyt_count <- function(q, date1, date2){
  r <- GET(base_url, query=list(q=q,
                                "api-key"=apikey,
                                "begin_date"=date1,
                                "end_date"=date2))

  json <- content(r, "parsed")
  return(json$response$meta$hits)
}

# Llamamos a la función en un periodo dado
nyt_count(q="Green Deal", date1=20180101, date2=20210101)

# Podemos hacerlo más complejo aún
nyt_years_count <- function(q, yearinit, yearend){
  # sequence of years to loop over
  years <- seq(yearinit, yearend)
  counts <- rep(NA, length(years))
  # loop over periods
  for (i in 1:length(years)){
    # information message to track progress
    message(years[i])
    # retrieve count
    counts[i] <- nyt_count(q=q, date1=paste0(years[i], "0101"),
                          date2=paste0(years[i], "1231"))
  }
  return(counts)
}

# vamos a ver qué sucede...
nyt_years_count(q="Green Deal", yearinit=2018, yearend=2021)

# ... estamos haciendo llamadas muy rápido... le metemos un Sys.sleep
nyt_count <- function(q, date1, date2){
  r <- GET(base_url, query=list(q=q,
                                "api-key"=apikey,
                                "begin_date"=date1,
                                "end_date"=date2))

  json <- content(r, "parsed")

```

```

## if there is no response
while (r$status_code!=200){
  Sys.sleep(2) # wait a couple of seconds
  # try again:
  r <- GET(base_url, query=list(q=q,
                                "api-key"=apikey,
                                "begin_date"=date1,
                                "end_date"=date2))

  json <- content(r, "parsed")
}
return(json$response$meta$hits)
}

# And let's see if this does the trick...
counts <- nyt_years_count(q="Green Deal", yearinit=2018,
                          yearend=2021)

plot(2018:2021, counts, type="l", main="Menciones de Green Deal por
año en NYT",
      xlab="Año", ylab="Número de artículos")

# Hacemos la función para cualquier intervalo de fechas
nyt_dates_count <- function(q, init, end, by){
  # sequence of dates to loop over
  dates <- seq(from=init, to=end, by=by)
  dates <- format(dates, "%Y%m%d") # changing format to match NYT
API format
  counts <- rep(NA, length(dates)-1)
  # loop over periods
  for (i in 1:(length(dates)-1)){ ## note the -1 here
    # information message to track progress
    message(dates[i])
    # retrieve count
    counts[i] <- nyt_count(q=q, date1=dates[i],
                          date2=dates[i+1])
  }
  # improving this as well so that it returns a data frame
  df <- data.frame(date = as.Date(dates[-length(dates)]),
                  format="%Y%m%d"), count = counts)
}

```



```

    return(df)
}

# Y la usamos
counts <- nyt_dates_count(q="Green Deal", init =
as.Date("2018/01/01"), end = as.Date("2021/06/01"), by="month")
plot(counts$date, counts$count, type="l", main="Menciones de Green
Deal en el intervalo, por mes",
      xlab="Mes", ylab="Número de artículos")

```

### **8.2. Código utilizado para la descarga de las cotizaciones de Repsol y para la elaboración gráfico que muestra su evolución.**

```

install.packages("quantmod")
library(quantmod)
getSymbols("REP.MC",src="yahoo")
head(cotizaciones)
tail(cotizaciones)
chartSeries(REP.MC, subset='2015::2020')

```

### **8.3. Código utilizado en *Python* para la descarga de los *tweets* sobre Repsol y Sostenibilidad en el período que va de enero de 2015 hasta diciembre de 2020.**

```

# -*- coding: utf-8 -*-
"""
Created on Wed June 2 12:05:10 2021

@author: Carlos Varona Vidaurrezaga
"""

from platform import python_version
print(python_version())
!pip install snsrape
#!pip install vaderSentiment

```

```

import pandas as pd
import numpy as np
import sns scrape.modules.twitter as sntwitter
import csv
import matplotlib.pyplot as plt

# Set maximum tweets to pull
maxTweets = 20000
# Set what keywords you want your twitter scraper to pull
keyword = 'Repsol Sostenibilidad'
#Open/create a file to append data to
csvFile = open('pfe_tweets_resultREPSOL.csv', 'a', newline='',
encoding='utf8')
#Use csv writer
csvWriter = csv.writer(csvFile)
csvWriter.writerow(['id','date','tweet',])

# Write tweets into the csv file
for i,tweet in enumerate(sntwitter.TwitterSearchScrapper(keyword + '
lang:es since:2015-01-01 until:2020-12-30 -filter:links -
filter:replies').get_items()):
    if i > maxTweets :
        break
    csvWriter.writerow([tweet.id, tweet.date, tweet.content])
csvFile.close()

```

#### **8.4. Códigos utilizados en *RStudio* para la realización del análisis de sentimiento de los *tweets* descargados previamente.**

Para el análisis de sentimientos se han utilizado tres *scripts* diferentes:

En primer lugar, para la obtención de las Ilustraciones 24, 25 y 26, se utilizó el siguiente código:

```

# Primero, limpiamos la memoria de trabajo
rm(list = ls())

```

```

#Paquetes necesarios
library(tidyverse)
library(rvest)
library(tidytext)
library(udpipe)
library(data.table)
library(BTM)
library(widyr)
library(igraph)
library(ggraph)
library(quanteda)
library(topicmodels)
library(scales)

#Leemos las opiniones sobre la gestión de Repsol sobre la
sostenibilidad
datos<-read.csv("pfe_tweets_result.csv")
view(datos)

#Quitamos lo que no nos hace falta, quedándonos con solo los
comentarios
datos<-datos[c(3)]
view(datos)
str(datos)
#Lo convertimos en un data frame
datos<-as.data.frame(datos)

# Eliminamos las palabras innecesarias para analizar los comentarios
lexiconSW<-stopwords("es")
lexiconSW<-
append(lexiconSW,c("Repsol","sostenibilidad","sostenible","Director
","@repsol","luis","dueñas","cabra","general","#repsol"))
lexiconSW<-as.data.frame(lexiconSW)
names(lexiconSW)<-"palabra"
lexiconSW$palabra<-as.character(lexiconSW$palabra)

#Preparamos el tipo de los datos y le damos su estructura
datos$comentarios<-as.character(datos$tweet)
df <- tibble::rowid_to_column(datos, "ID")

```

```

#Eliminamos las stopwords
lista_palabras <- df %>%
  dplyr::distinct(comentarios, .keep_all = TRUE) %>%
  unnest_tokens(palabra, comentarios, drop = FALSE) %>%
  distinct(ID, palabra, .keep_all = TRUE) %>%
  anti_join(lexiconSW) %>%
  filter(str_detect(palabra, "[^\\d]")) %>%
  group_by(palabra) %>%
  dplyr::mutate(word_total = n()) %>%
  ungroup()

#creamos un corpus para el procesamiento semántico+

corpus <- corpus(datos$comentarios)
cdfm <- dfm(cdfm, corpus,
remove=c(stopwords("spanish"), "Repsol", "sostenibilidad", "sostenible",
"Director", "@repsol", "luis", "dueñas", "cabra", "general", "#repsol",
"@fundacionrepsol", "#sostenibilidad"),
verbose=TRUE, remove_punct=TRUE, remove_numbers=TRUE)
cdfm <- dfm_trim(cdfm, min_docfreq = 2, verbose=TRUE)

#buscamos un gráfico que nos muestre las palabras más utilizado
textplot_wordcloud(cdfm,
rot.per=0,
scale=c(5.5, 1.25),
max.words=500,
color = rev(RColorBrewer::brewer.pal(10,
"RdBu"))))

#Llevamos a cabo un analisis de bigramas
pares_palabras <- df %>%
  unnest_tokens(palabra, comentarios) %>%
  anti_join(lexiconSW) %>%
  pairwise_count(palabra, ID, sort = TRUE, upper = FALSE)

#Ahora hacemos el bigrama graficamente
set.seed(1234)
pares_palabras %>%

```

```

filter(n >= 20) %>%
graph_from_data_frame() %>%
ggraph(layout = "fr") +
geom_edge_link(aes(edge_alpha = n, edge_width = n), edge_colour =
"cyan4") +
geom_node_point(size = 5) +
geom_node_text(aes(label = name), repel = TRUE,
               point.padding = unit(0.2, "lines")) +
ggtitle('Bigram Graph: combination of 2 keywords')

#hacemos analisis cluster
corpus <- corpus(datos$comentarios)
cdfm <- dfm(corpus, remove=c(stopwords("spanish"),

"Repsol", "sostenibilidad", "sostenible", "Director", "@repsol", "luis",
"dueñas", "cabra", "general", "#repsol"),
            verbose=TRUE, remove_punct=TRUE, remove_numbers=TRUE)
cdfm <- dfm_trim(cdfm, min_docfreq = 2, verbose=TRUE)

#topicmodels
dtm <- convert(cdfm, to="topicmodels")

#he tratado de encontrar el n° óptimo de clusters:
library(lattice)
library(robustbase)
library(cvTools)
cvLDA <- function(Ntopics, dtm, K=10) {
  folds<-cvFolds(nrow(dtm), K, 1)
  perplex <- rep(NA, K)
  llk <- rep(NA, K)
  for(i in unique(folds$which)) {
    cat(i, " ")
    which.test <- folds$subsets[folds$which==i]
    which.train <- {1:nrow(dtm)}[-which.test]
    dtm.train <- dtm[which.train,]
    dtm.test <- dtm[which.test,]
    lda.fit <- LDA(dtm.train, k=Ntopics, method="Gibbs",
                  control=list(verbose=50L, iter=100))
    perplex[i] <- perplexity(lda.fit, dtm.test)
  }
}

```

```

    llk[i] <- logLik(lda.fit)
  }
  return(list(K=Ntopics,perplexity=perplex,logLik=llk))
}

numeroTopicsOptimo<-function(dtm){
  K <- c(5,10,20, 30, 40, 50, 60, 70, 80)
  results <- list()

  i = 1
  for (k in K){
    cat("\n\n\n#####\n ", k, "topics", "\n")
    res <- cvLDA(k, dtm)
    results[[i]] <- res
    i = i + 1
  }

  ## plot
  df <- data.frame(
    k = rep(K, each=10),
    perp = unlist(lapply(results, '[' , 'perplexity')),
    loglk = unlist(lapply(results, '[' , 'logLik')),
    stringsAsFactors=F)

  min(df$perp)
  df$ratio_perp <- df$perp / max(df$perp)
  df$ratio_lk <- df$loglk / min(df$loglk)

  df <- data.frame(cbind(
    aggregate(df$ratio_perp, by=list(df$k), FUN=mean),
    aggregate(df$ratio_perp, by=list(df$k), FUN=sd)$x,
    aggregate(df$ratio_lk, by=list(df$k), FUN=mean)$x,
    aggregate(df$ratio_lk, by=list(df$k), FUN=sd)$x),
    stringsAsFactors=F)
  names(df) <- c("k", "ratio_perp", "sd_perp", "ratio_lk", "sd_lk")
  library(reshape)
  pd <- melt(df[,c("k","ratio_perp", "ratio_lk")], id.vars="k")
  pd2 <- melt(df[,c("k","sd_perp", "sd_lk")], id.vars="k")
  pd$sd <- pd2$value

```

```

levels(pd$variable) <- c("Perplexity", "LogLikelihood")

library(ggplot2)
library(grid)

p <- ggplot(pd, aes(x=k, y=value, linetype=variable))
pq <- p + geom_line() + geom_point(aes(shape=variable),
                                   fill="white", shape=21,
size=1.40) +
  geom_errorbar(aes(ymax=value+sd, ymin=value-sd), width=4) +
  scale_y_continuous("Ratio wrt worst value") +
  scale_x_continuous("Number of topics",
                    breaks=K) +
  theme_bw()
pq
return(pq)
}

numeroTopicsOptimo(dtm)
#he seleccionado el numero de clusters en 20
lda <- LDA(dtm, k = 20, method = "Gibbs",
          control = list(verbose=25L, seed = 123, burnin = 100, iter
= 500))
#vemos las palabras mas representativas
terms(lda)
#las top 10 ahora
trms <- t(terms(lda, k=10))

#por separado
trms[6,]
trms[1,]
trms[9,]
trms[13,]
trms[4,]
trms[9,]
trms[19,]

#buscamos sacar el peso de cada palabra con BETA
terminosTopic <- tidy(lda, matrix = "beta")
#peso semantico en cada topic

```

```

# Sacamos los datos de cada topic
terminosTopic <- tidy(lda, matrix = "beta")
terminosTopic
#primeros terminos de cada topic
top_terms <- terminosTopic %>%
  group_by(topic) %>%
  top_n(8, beta) %>%
  ungroup() %>%
  arrange(topic, -beta)
#lo vemos:
top_terms %>%
  mutate(term = reorder(term, beta)) %>%
  ggplot(aes(term, beta, fill = factor(topic))) +
  geom_col(show.legend = FALSE) +
  facet_wrap(~ topic, scales = "free") +
  coord_flip()+
  labs(x = NULL, y = "Palabras más importantes por tema",
       title = paste0("Temas y sus palabras más descriptivas"))

```

En segundo lugar, para el análisis de sentimientos propiamente dicho, se han utilizado dos códigos.

El primero que vamos a adjuntar es el utilizado para las Ilustraciones 27 y 28, mientras que el segundo de ellos es el usado para el resto de ilustraciones referentes al análisis de sentimientos.

```

tuits <- read.csv("pfe_tweets_result.csv")

#Instalación de paquetes que se utilizarán

install.packages("SnowballC")

install.packages("tm")

install.packages("twitter")

install.packages("syuzhet")

```



```

library(syuzhet)

#Verificando cantidad de tweets importados

n.tweet <- length(tuits)

n.tweet

#Convirtiendo los tweets en un data frame

tweets.df <- twListToDF(tuits)

head(tweets.df)

#Quitando los links en los tweets

tuits <- gsub("http.*","",tweets.df$text)

tuits <- gsub("https.*","",tuits)

#Quitando los hashtags y usuarios en los tweets

tuits <- gsub("#\\w+","",tuits)

tuits <- gsub("@\\w+","",tuits)

#Quitando los signos de puntuación, números y textos con números

tuits <- gsub("[[:punct:]]","",tuits)

tuits <- gsub("\\w*[0-9]+\\w*\\s*", "",tuits)

#Transformamos la base de textos importados en un vector para

#poder utilizar la función get_nrc_sentiment

palabra.df <- as.vector(tuits)

#Aplicamos la función indicando el vector y el idioma y creamos

#un nuevo data frame llamado emocion.df

emocion.df <- get_nrc_sentiment(char_v = palabra.df, language =
"spanish")

```

```

#Unimos emocion.df con el vector tweets.df para ver como
#trabajó la función get_nrc_sentiment cada uno de los tweets
emocion.df2 <- cbind(tuits, emocion.df)

head(emocion.df2)

#Creamos un data frame en el cual las filas serán las emociones
#y las columnas los puntajes totales

#Empezamos transponiendo emocion.df
emocion.df3 <- data.frame(t(emocion.df))

#Sumamos los puntajes de cada uno de los tweets para cada emocion
emocion.df3 <- data.frame(rowSums(emocion.df3))

#Nombramos la columna de puntajes como cuenta
names(emocion.df3)[1] <- "cuenta"

#Dado que las emociones son los nombres de las filas y no una variable
#transformamos el data frame para incluirlas dentro

emocion.df3 <- cbind("sentimiento" = rownames(emocion.df3),
emocion.df3)

#Quitamos el nombre de las filas
rownames(emocion.df3) <- NULL

#Verificamos el data frame

print(emocion.df3)

library(ggplot2)

```

```
#Primer gráfico: se detallaran las 8 emociones con sus puntajes respectivos
```

```
sentimientos1 <- ggplot(emocion.df3[1:8,],  
  
                        aes(x = sentimiento,  
  
                            y = cuenta, fill = sentimiento)) +  
  
  geom_bar(stat = "identity") +  
  
  labs(title = "Emociones con mayor frecuencia en los tweets sobre  
Repsol y Sostenibilidad",  
  
        x = "Sentimiento", y = "Frecuencia") +  
  
  geom_text(aes(label = cuenta),  
  
            vjust = 1.5, color = "black",  
  
            size = 5) +  
  
  theme(plot.title = element_text(hjust = 0.5),  
  
        axis.text = element_text(size=12),  
  
        axis.title = element_text(size=14,face = "bold"),  
  
        title = element_text(size=20,face = "bold"),  
  
        legend.position = "none")  
  
print(sentimientos1)
```

```
#Segundo gráfico: se detallan los puntajes para las valoraciones
```

```
#positiva y negativa
```

```
sentimientos2 <- ggplot(emocion.df3[9:10,],  
  
                        aes(x = sentimiento,  
  
                            y = cuenta, fill = sentimiento)) +
```

```

geom_bar(stat = "identity") +

labs(title = "Sentimiento predominante",

      x = "Sentimiento", y = "Frecuencia") +

geom_text(aes(label = cuenta),

          vjust = 1.5, color = "black",

          size = 5) +

theme(plot.title = element_text(hjust = 0.5),

      axis.text = element_text(size=12),

      axis.title = element_text(size=14,face = "bold"),

      title = element_text(size=20,face = "bold"),

      legend.position = "none")

print(sentimientos2)

```

Para el resto del análisis de sentimiento, este es el código utilizado:

```

# Instalamos los paquetes:

install.packages("syuzhet")

install.packages("RColorBrewer")

install.packages("wordcloud")

install.packages("tm")

install.packages("devtools")

devtools::install_github("ankitrohatgi/digitizeR")

# Cargamos los paquetes

```

```

library(syuzhet)

library(RColorBrewer)

library(wordcloud)

library(tm)

library("devtools")

library('digitizeR')

library("tidyverse")

#cargamos los tweets

tuits <- read.csv("pfe_tweets_result.csv")

#funciones:

texto_palabras <- get_tokens(tuits)

head(texto_palabras)

length(texto_palabras)

sentimientos_df <- get_nrc_sentiment(texto_palabras, lang="spanish")

summary(sentimientos_df)

barplot(

  colSums(prop.table(sentimientos_df[, 1:8])),

  space = 0.2,

  horiz = FALSE,

  las = 1,

  cex.names = 0.7,

  col = brewer.pal(n = 8, name = "Set3"),

```

```

main = "Tweets sobre Repsol y la Sostenibilidad",

sub = "Análisis realizado por Carlos Varona Vidaurrazaga",

xlab="emociones", ylab = NULL)

barplot(colSums(prop.table(sentimientos_df[, 1:8])))

palabras_tristeza <- texto_palabras[sentimientos_df$sadness > 0]

palabras_confianza <- texto_palabras[sentimientos_df$trust > 0]

palabras_confianza_orden <- sort(table(unlist(palabras_confianza)),
decreasing = TRUE)

head(palabras_confianza_orden, n = 12)

length(palabras_confianza_orden)

head(palabras_confianza_orden, n = 12)

length(palabras_tristeza_orden)

nube_emociones_vector <- c(

  paste(texto_palabras[sentimientos_df$sadness > 0], collapse = " "),

  paste(texto_palabras[sentimientos_df$joy > 0], collapse = " "),

  paste(texto_palabras[sentimientos_df$anger > 0], collapse = " "),

  paste(texto_palabras[sentimientos_df$fear > 0], collapse = " "))

nube_corpus <- Corpus(VectorSource(nube_emociones_vector))

nube_tdm <- TermDocumentMatrix(nube_corpus)

nube_tdm <- as.matrix(nube_tdm)

head(nube_tdm)

```

```

colnames(nube_tdm) <- c('tristeza', 'felicidad', 'enfado',
'confianza', "miedo", "anticipación", "alegría", "sorpresa")

head(nube_tdm)

set.seed(757) # puede ser cualquier número

comparison.cloud(nube_tdm, random.order = FALSE,

                 colors = c("green", "red", "orange", "blue"),

                 title.size = 1, max.words = 50, scale = c(2.5, 1),
rot.per = 0.4)

sentimientos_valencia <- (sentimientos_df$negative *-1) +
sentimientos_df$positive

simple_plot(sentimientos_valencia)

```

### 8.5. Código utilizado de *RStudio* para obtener le gráfico con la evolución del sentimiento de los *tweets*

Por último, se muestra a continuación el código utilizado para extraer el sentimiento promedio semanal de los *tweets* sobre Repsol y Sostenibilidad.

```

library(rtweet)
library(tidytext)
library(dplyr)
library(tidyverse)
library(wordcloud)
library(reshape2)
library(RColorBrewer)

tuits <- read.csv("pfe_tweets_result.csv")
download.file("https://raw.githubusercontent.com/jboscomendoza/rpub
s/master/sentimientos_afinn/lexico_afinn.en.es.csv",
             "lexico_afinn.en.es.csv")

```

```

afinn <- read.csv("lexico_afinn.en.es.csv", stringsAsFactors = F,
fileEncoding = "latin1") %>%
  tbl_df()
tuits_afinn <-
  tuits %>%
  unnest_tokens(input = "tweet", output = "Palabra") %>%
  inner_join(afinn, ., by = "Palabra") %>%
  mutate(Tipo = ifelse(Puntuacion > 0, "Positiva", "Negativa"))

tuits <-
  tuits_afinn %>%
  group_by(date) %>%
  summarise(Puntuacion_tuit = mean(Puntuacion)) %>%
  left_join(tuits, ., by = "date") %>%
  mutate(Puntuacion_tuit = ifelse(is.na(Puntuacion_tuit), 0,
Puntuacion_tuit))
tuits_afinn %>%
  count(date)
library(purrr)

tuits_afinn %>%
  ggplot() +
  aes(date, Puntuacion, color = Puntuacion) +
  geom_smooth(method = "loess", fill = NA)
Sentimiento <- ts(tuits_afinn$Puntuacion, start = 2015, frequency =
52)
plot(Sentimiento)

```



