



Facultad de Ciencias Económicas y Empresariales
ICADE

**Análisis de la discrepancia entre los múltiplos de
valoración de empresas PER y EV/EBITDA: un
enfoque por industria**

Autor: Carlos de Abajo Sáenz-Diez

Director: Carlos Bellón Núñez-Mera

MADRID | Junio 2024

RESUMEN

La búsqueda del valor intrínseco de los activos financieros es una tarea clave para los profesionales del mercado, quienes, entre otras técnicas, emplean múltiplos de valoración como el PER o el EV/EBITDA. Estos ratios ayudan a evaluar si el precio de una acción es adecuado, sobrevalorado o infravalorado en comparación con empresas similares, influyendo en decisiones de inversión y ajustes en las carteras. Sin embargo, para que esta metodología sea efectiva, es esencial interpretar correctamente estos múltiplos considerando los fundamentos y la estrategia de la empresa.

La interpretación se complica cuando los ratios presentan contradicciones: un múltiplo puede indicar que una acción está sobrevalorada mientras otro sugiere lo contrario. Este TFG se centra en la identificación de estas discrepancias, que hacen que una acción parezca "barata según el P/E pero cara según el EV/EBITDA". Utilizando la base de datos compustat y el lenguaje de programación python, se calculan los múltiplos de valoración diarios para 5.679 empresas durante el periodo 2010-2023. Tras clasificar estos múltiplos en cuartiles, se reconoce la existencia de discrepancias y se investiga si tales discrepancias ocurren con mayor frecuencia en ciertas industrias, así como los posibles motivos que las justifiquen.

Palabras clave: múltiplos de valoración, PER, EV/EBITDA, discrepancia, valoración de acciones

ABSTRACT

The search for the intrinsic value of financial assets is a key task for market professionals, who use valuation multiples such as PER or EV/EBITDA among other techniques. These ratios help evaluate whether the price of a stock is appropriate, overvalued, or undervalued compared to similar companies, influencing investment decisions and portfolio adjustments. However, for this methodology to be effective, it is essential to interpret these multiples correctly, considering the company's fundamentals and strategy.

Interpretation becomes complicated when the ratios present contradictions: one multiple may indicate that a stock is overvalued while another suggests the opposite. This TFG focuses on identifying these discrepancies, which make a stock appear "cheap on P/E but expensive on EV/EBITDA." Using the compustat database and python programming language, daily valuation multiples are calculated for 5,679 companies during the period 2010-2023. After classifying these multiples into quartiles, the existence of discrepancies is recognized, and we

try to determine if such discrepancies occur more frequently in certain industries and the possible reasons behind them.

Key words: Valuation multiples, PER, EV/EBITDA, discrepancies, stock valuation

ÍNDICE

CAPÍTULO I: INTRODUCCIÓN.....	6
1.1 Interés del tema.....	6
1.2 Objetivos.....	8
1.3 Estructura.....	8
CAPÍTULO II: MARCO TEÓRICO. VALORACIÓN DE ACCIONES Y MÚLTIPLOS DE COTIZACIÓN.....	9
2.1 Concepto y breve evolución.....	9
2.2 El uso de múltiplos y la popularización del EV/EBITDA.....	11
2.3 Algunos estudios de valoración por múltiplos.....	18
CAPÍTULO III: METODOLOGÍA: CONSTRUCCIÓN DE LA MUESTRA Y CÁLCULO DE LOS MÚLTIPLOS.....	21
3.1 Bases de datos utilizadas.....	21
3.2 Clasificación por industria: códigos SIC.....	21
3.3 Definición de múltiplos y variables compustat utilizadas.....	25
a. Variables históricas TTM (Trailing Twelve Months).....	25
b. Definición de múltiplos y variables de compustat.....	26
3.4. Universo de compañías.....	31
3.5. Construcción de los múltiplos.....	32
3.6. Cálculo de las discrepancias.....	33
CAPÍTULO IV: RESULTADOS E INTERPRETACIÓN.....	35
4.1. Descripción de resultados.....	35
4.2. Interpretación de resultados. Análisis de las industrias.....	39
CAPÍTULO V: CONCLUSIONES.....	47
CAPÍTULO VI: BIBLIOGRAFÍA.....	52
CAPÍTULO VII: ANEXOS.....	57
7.1 Código Compustat Cleaning.....	57
7.2 Código Prices cleaning.....	63
7.3 Código Multiples.....	69
7.4 Código Cálculo discrepancias.....	72

ÍNDICE DE FIGURAS:

Figura 1: Enfoques de valoración más utilizados entre los encuestados en el Estudio de 2015 del Instituto CFA

Figura 2: Métodos de valoración preferidos

Figura 3: Múltiplo de mercado más utilizado entre los encuestados en el Estudio de 2015 del Instituto CFA

Figura 4: Principal múltiplo utilizado por el participante

Figura 5: Múltiplo preferido para cada industria

Figura 6: Referencia a “EBITDA” y a “EPS” en libros, 1960-2008

Figura 7: Popularidad de los diferentes múltiplos de valoración

Figura 8. Composición de los dígitos del código SIC

Figura 8: Distribución de discrepancias por industria

Figura 9: Discrepancia media por empresa e industria. Índice de discrepancia media

ÍNDICE DE TABLAS:

Tabla 1. Artículos académicos sobre múltiplos de valoración: definición de las variables principales

Tabla 2. Reconciliación del múltiplo EV/EBITDA y el PER

Tabla 4. Divisiones industriales según el SIC de un dígito

Tabla 4: Códigos SIC-dos dígitos

Tabla 5: Variables anuales y trimestrales de compustat para construir múltiplos

Tabla 6: Descripción de las discrepancias

Tabla 7: Industrias con mayor número de empresas

Tabla 8: Frecuencia de discrepancias por industria

Tabla 9: Porcentaje de discrepancias por industria

CAPÍTULO I: INTRODUCCIÓN

1.1 Interés del tema

La valoración relativa o valoración por múltiplos es uno de los métodos más extendidos en la práctica diaria de la industria financiera, utilizados para estimar el “precio objetivo” o “precio justo” al que debería cotizar una acción si el mercado fuera eficiente y reflejara de forma adecuada el valor intrínseco de los activos.

La hipótesis del mercado eficiente es uno de los principios fundamentales de las finanzas clásicas y de la teoría de la valoración de activos, y sostiene que los precios de los activos reflejan toda la información disponible en cualquier momento, de modo que los inversores no pueden lograr consistentemente rendimientos que sean superiores a los que corresponden al nivel de riesgo asumido (Fama, 1970). En mercados eficientes, los precios incorporan rápidamente la nueva información mediante la interacción de los inversores que reaccionan a las noticias comprando o vendiendo activos financieros (Fama, 1965; Grossman & Stiglitz, 1976).

Esta hipótesis ha sido objeto de debate desde su formulación en la década de 1970 sin un consenso en la literatura sobre su validez (Boettke, 2010). Son muchos los estudios sobre eficiencia de mercado que se han realizado a lo largo de los años (Jensen, 1978 o De Bondt & Thaler, 1985). Si bien la hipótesis de partida en los modelos de finanzas clásicas es la de mercados eficientes en grado de eficiencia fuerte, los trabajos empíricos han demostrado que el grado de eficiencia en realidad es, en el mejor de los casos, semi fuerte (Ball y Brown, 2013 o Fama et al, 1969), y en momentos puntuales, ni siquiera esa (Lo & MacKinlay, 1988 o Jegadeesh & Titman, 1993). Esta falta de eficiencia supone un acicate para los agentes del mercado bursátil, que intentarán batir al mercado buscando situaciones de *mispricing* que supongan oportunidades atractivas de inversión.

El trabajo del analista financiero consiste, por tanto, en descubrir si un activo financiero tiene un precio incorrecto (es decir, está infravalorado o sobrevalorado con respecto a su valor intrínseco o al precio de activos similares), suponiendo que en algún momento los mercados corregirán el error en el precio. La habilidad para identificar situaciones de *mispricing* y la capacidad de aprovecharlas producirían beneficios cuando los precios de los activos se aproximaran a su valor real o volvieran a su posición correcta en relación a activos comparables.

En este sentido, los múltiplos de valoración se utilizan a menudo como un primer filtro para establecer si el precio de una acción parece correcto o es “caro” o “barato” en relación a activos similares. Sin embargo, no hay que olvidar que una empresa puede cotizar de forma continuada por encima (o por debajo) de los niveles medios de sus comparables y no por ello su precio es incorrecto, sino que su calidad, ventaja competitiva, generación de flujos de caja o riesgo, justifican esa prima (o descuento) respecto al resto de la industria¹.

La interpretación resulta aún más compleja cuando los ratios de valoración generan una discrepancia: según un múltiplo el precio de la acción estaría por encima de la media de sus comparables, pero según otro estaría por debajo. Es decir, la acción puede estar “*cheap on P/E but expensive on EV/EBITDA*”² o a la inversa. Este Trabajo Fin de Grado (TFG) calcula la presencia diaria de estas discrepancias en una muestra de 5.679 compañías durante un periodo de casi 13 años (desde el 1 de abril de 2010 al 29 de diciembre de 2023) y analiza estas discrepancias así como algunos de los posibles motivos detrás de las mismas.

La mayoría de los trabajos sobre valoración por múltiplos descansan en la hipótesis de mercados eficientes, de la que se deriva que los precios de mercado son siempre correctos. Junto a ellos, se encuentran otros estudios (e. g. Basu, 1977 o Stattman, 1980) en los que el papel de los múltiplos es identificar acciones incorrectamente valoradas por el mercado y que pueden proporcionar rendimientos “anormales” a sus inversores. Podría decirse que este TFG se encuadra en esta línea de investigación: construir múltiplos con el precio de acciones cotizadas tratando de determinar si la acción está “cara” o “barata” en relación a sus comparables, y si por tanto podría ser una buena oportunidad de inversión. Concretamente, intentaremos analizar qué ocurre cuando los múltiplos de una empresa arrojan resultados contradictorios, siendo “barata” según el ratio PER y “cara” según el EV/EBITDA o a la inversa.

Este TFG se engarza en una línea de investigación liderada por el Profesor Carlos Bellón, y se apoya en parte en trabajos anteriores como el de Urbano García (2023) o el de Álvarez-Sala (2022).

¹ La valoración relativa descansa en la premisa de comparabilidad entre activos (empresas en este caso), pero esta comparabilidad es en realidad muy imperfecta, lo que supone la principal desventaja conceptual de este enfoque. “*There is no such thing as two similar companies*” es el argumento que se esgrime con frecuencia para interpretar con mucha cautela las conclusiones que se extraen de la posición relativa de un múltiplo respecto a la media de su industria.

² Jullens (2012).

1.2 Objetivos

El propósito general de este TFG es dar un paso más en el análisis y comprensión de las discrepancias existentes en los múltiplos de valoración contruidos con precios de las acciones.

Para ello, es necesario en primer lugar construir una muestra de empresas, ordenadas por industrias, de las que se obtengan datos contables y precios diarios, que son la materia prima a partir de la cual se construirán los múltiplos. Posteriormente, se calcularán los múltiplos de valoración diarios para cada una de las 5.679 empresas durante el periodo 2010-2023. Se ordenarán en cuartiles dichos múltiplos, dentro de cada industria, para reconocer la existencia de discrepancias. Por último, se tratará de identificar si tales discrepancias suceden con más frecuencia en ciertas industrias y los posibles motivos que las justifiquen.

Concretando, los objetivos de este TFG serían:

1. Utilizar la base de datos compustat para construir los múltiplos de valoración PER e EV/EBITDA y calcular las posibles discrepancias.
2. Describir las discrepancias encontradas
3. Reflexionar sobre la mayor o menor frecuencia de discrepancias por industria, tratando de ligar la presencia de las mismas a las características del negocio

1.3 Estructura

El resto del TFG se estructura de la siguiente manera: el capítulo II explica el marco teórico que da contexto a este trabajo, explicando la utilidad y popularidad de los múltiplos de valoración, tanto en la práctica de los mercados como en trabajos académicos. El capítulo III se canta en la metodología de trabajo, desde la descripción de la base de datos compustat o la agrupación de empresas en industrias definidas con el código SICa la definición de los múltiplos y de las variables que los componen, para terminar explicando cómo se calculan los múltiplos y las discrepancias, utilizando para todo ello programación con python. El capítulo IV expone los resultados y la interpretación de los mismos, y el capítulo V recoge las principales conclusiones del trabajo. Por último, las referencias bibliográficas pueden consultarse en el capítulo VI y los códigos que construyen los *datasets* y calculan las discrepancias en el capítulo VII de anexos.

CAPÍTULO II: MARCO TEÓRICO. VALORACIÓN DE ACCIONES Y MÚLTIPLOS DE COTIZACIÓN

2.1 Concepto y breve evolución

Aswath Damodaran, un conocido profesor de finanzas en la Stern School of Business de la Universidad de Nueva York, define “valoración” como el proceso de estimar el valor de un activo, negocio o inversión. Según Damodaran, la valoración es una combinación de arte y ciencia que implica hacer estimaciones razonables sobre los flujos de caja futuros de un activo, su potencial de crecimiento y su riesgo, para después obtener su valor en el momento actual. Este proceso requiere una comprensión profunda de las características del activo, la industria en la que opera y las condiciones generales del mercado. La valoración se trata, en última instancia, de hacer juicios y conjeturas coherentes y sensatas, con el objetivo de llegar a una estimación razonable del valor (Damodaran, 2012).

Como recuerda Mauboussin (2018), existe una rica literatura financiera sobre valoración, que es sencilla conceptualmente pero a menudo complicada de aplicar en la práctica. Existe un acuerdo general sobre que un negocio vale el valor actual de sus flujos de caja libres futuros, pero pronosticar la magnitud, el momento y el riesgo de estos flujos presenta enormes dificultades. Por este motivo, es muy habitual en la práctica diaria de los mercados recurrir a métodos simplificados o “atajos” en la valoración de acciones utilizando múltiplos como el ratio precio-beneficio (PER por las siglas en inglés de *price-earnings ratio*) o el EV/EBITDA.

Las técnicas de valoración de acciones han ido evolucionando a lo largo del tiempo. Hace cien años, en la bolsa norteamericana, los inversores valoraban las acciones basándose en métricas como la rentabilidad por dividendo y los múltiplos precio-beneficio (*Price Earnings ratio* o PER) y precio-valor contable (*Price to Book Value* o P/BV). Benjamin Graham y David Dodd publicaron en 1934 uno de los primeros textos clásicos en la materia: *Security Analysis*, en el que afirman que la valoración de acciones ordinarias se basaba fundamentalmente en la tasa de dividendo, los beneficios y el valor de los activos. Poco después, en 1938, John Burr Williams formalizó el modelo de descuento de dividendos en otro texto seminal: *The Theory of Investment Value*. En 1959 Myron Gordon estableció el modelo de crecimiento de dividendos.

Los modelos de valoración que conocemos en la actualidad se construyen a partir del artículo publicado en 1961 por los profesores Merton Miller y Franco Modigliani *Dividend*

Policy, Growth, and the Valuation of Shares. En este trabajo se plantearon una pregunta fundamental: ¿qué descuenta el mercado de valores? Consideraron cuatro posibilidades: flujo de caja, beneficios actuales más oportunidades futuras de creación de valor, la corriente de dividendos y la corriente de beneficios. Su conclusión fue sorprendente: demostraron que correctamente, considerados todos ellos convergían en el mismo modelo. Así, el valor de una acción es el valor actual de los flujos de caja libres futuros, modelo que se conoce popularmente como DCF y que se estudia actualmente en escuelas de negocios de todo el mundo.

La puesta en práctica de este modelo no carece de retos y dificultades. La razón principal es que a pesar de que el modelo de flujos de caja descontados es analíticamente sólido, exige una serie de juicios de valor para aplicarse a la realidad. El resultado del análisis dependerá de la calidad de las hipótesis y estimaciones de la información utilizada. Platt et al (2010) tratan de comprobar empíricamente si lo cercano que está el *enterprise value* al valor actual de los flujos de caja una vez efectivamente se producen (y no a las estimaciones). Aunque este desafío puede abordarse de forma adecuada, muchos profesionales de la industria financiera evitan por completo un modelo DCF y en su lugar utilizan la valoración por múltiplos, más sencilla en apariencia y con menos requisitos de información para su implementación.

Es crucial reconocer que los múltiplos no son otra cosa que un resumen del proceso de valoración completo. Son un “atajo” que debería llevar al mismo resultado si son correctamente contruidos y utilizados. En este sentido, Liu et al (2001) afirman que aunque la valoración por múltiplos evita proyecciones explícitas y descuentos de los flujos de caja, descansa en los mismos principios que enfoques más completos como el DCF: el valor está directamente relacionado con los pagos futuros e inversamente relacionado con el riesgo. Los múltiplos pueden sustituir de forma efectiva a los modelos más detallados porque captan de forma adecuada la esencia de la valoración. Como afirman Koller et al (2015), mientras que el DCF es el método más correcto y flexible para valorar empresas, utilizar un enfoque de valoración relativa, puede proporcionar información y ayudar a resumir y comprobar el ejercicio de valoración. Sin embargo, los mismos autores recuerdan que a menudo los múltiplos se utilizan de una forma simplista y superficial, produciendo conclusiones erróneas. Para aplicar correctamente los múltiplos de valoración, estos mismos autores indican que es imprescindible elegir las compañías similares correctamente, así como desmenuzar e interpretar correctamente los estados financieros de las empresas para estar seguros de que los valores de los ratios son verdaderamente comparables (Koller et al, 2015).

2.2 El uso de múltiplos y la popularización del EV/EBITDA

Existen numerosos trabajos, a menudo basados en encuestas, estudian el uso de diferentes métodos de valoración por parte de agentes de mercado y profesionales de la industria financiera, como Block (2010), Bancel y Mittoo (2014), Fabozzi et al (2017), Nyborg y Mukhlynina (2016) o Pinto et al (2019)³. Las conclusiones de estos trabajos muestran rasgos comunes: El descuento de flujos de caja y los múltiplos de valoración son las metodologías más utilizadas en la práctica de los mercados, como se desprende de las figuras 1 y 2:

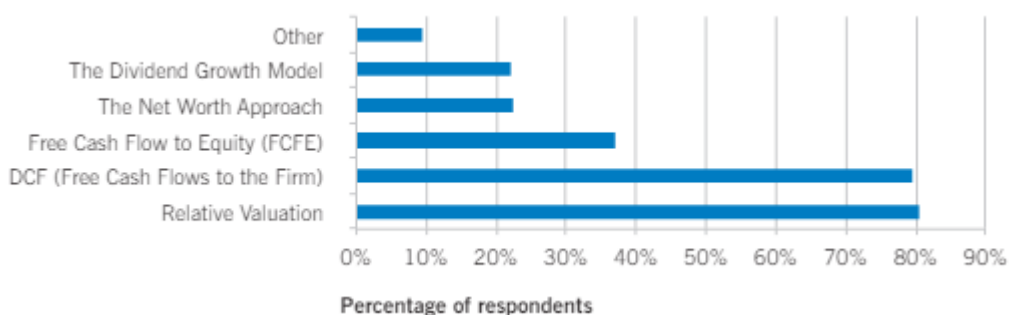
Figura 1: Enfoques de valoración más utilizados entre los encuestados en el Estudio de 2015 del Instituto CFA

Enfoques de valoración: Clasificación mundial. Al evaluar valores de renta variable individuales, ¿cuál de los siguientes enfoques de valoración utiliza? N=1980	Porcentaje de respuestas	Porcentaje de casos en los que el encuestado utiliza cada uno de los enfoques
Un enfoque de múltiplos de mercado	92.8	68.6
Un enfoque basado en el valor actual descontado	78.8	59.5
Enfoque basado en los activos	61.4	36.8
Enfoque de opciones (reales)	5.0	20.7
Otros	12.7	58.1

Fuente: Fabozzi et al. (2017)

³ Fabozzi et al (2017) y Pinto et al (2019) se basan en estudios y encuestas realizadas por el Instituto CFA.

Figura 2: Métodos de valoración preferidos



Fuente: Bancel y Mittoo (2014)

En cuanto a los múltiplos de valoración preferidos por los profesionales financieros, las distintas encuestas apuntan a que los dos múltiplos más populares son la relación precio-beneficio (ratio PER) y EV/EBITDA, como se refleja en la figura 3:

Figura 3: Múltiplo de mercado más utilizado entre los encuestados en el Estudio de 2015 del Instituto CFA

Enfoque de múltiplos de mercado: cuando utiliza un enfoque de múltiplos de mercado, ¿cuál de los siguientes ratios utiliza? N = 1,765	Porcentaje de respuestas	Porcentaje de casos en los que el encuestado utiliza cada uno de los enfoques
Rentabilidad del dividendo o precio del dividendo	35.5	44.3
Enterprise value o múltiplos de valor de empresa (EV/EBITDA o EV/beneficio de explotación)	76.7	61.1
Precio/valor contable, valor contable/mercado	59.0	44.8
Precio / Flujo de caja	57.2	54.6
Precio / ventas	40.3	45.7
Precio / beneficios	88.1	67.2
Otros	11.6	58.5

Fuente: Fabozzi et al. (2017)

Para estimar el precio de una acción, el múltiplo más utilizado tradicionalmente era el PER, que consiste en dividir el valor del *Equity* de la empresa (su capitalización bursátil si cotiza) entre el Beneficio Neto. Compañías en la misma industria, expuestas a los mismos riesgos, deberían cotizar con múltiplos similares (Koller et al, 2015).

A pesar de que, como se explicará posteriormente, los múltiplos basados en el EV (*Enterprise Value* o valor del negocio) pueden considerarse conceptualmente superiores, el PER es el múltiplo de valoración más conocido y utilizado. Según Block (2010), de los 10 manuales sobre inversiones más populares, solo uno menciona el múltiplo EV/EBITDA mientras que el PER se explica de forma detallada (entre cinco y siete páginas) en todos ellos. Este autor realiza una encuesta que respondieron 1.209 analistas financieros acerca del uso de múltiplos de valoración, recogiéndose las respuestas en la figura 4:

Figura 4: Principal múltiplo utilizado por el participante

Multiplier	Number	Percent
1. Price/earnings	504	41.7
2. EV/EBITDA	438	36.2
3. Price/book value	122	10.1
4. PEG (P/E/growth)	60	5.0
5. Forward price/earnings	52	4.2
6. Price/sales per share	15	1.3
7. Price/cash earnings	12	1.0
8. EV/sales	6	0.5
	1,209	100.0

Fuente: Block (2010)

La popularidad del PER es debida en parte a su simplicidad y también a que es el múltiplo más “universal”, pudiendo calcularse para cualquier tipo de empresas, industrias y mercados: los múltiplos de EV no son fácilmente aplicable a bancos y entidades financieras, por poner un ejemplo, pero sí es posible calcular el PER de dichas empresas, tal y como se observa en la figura 5:

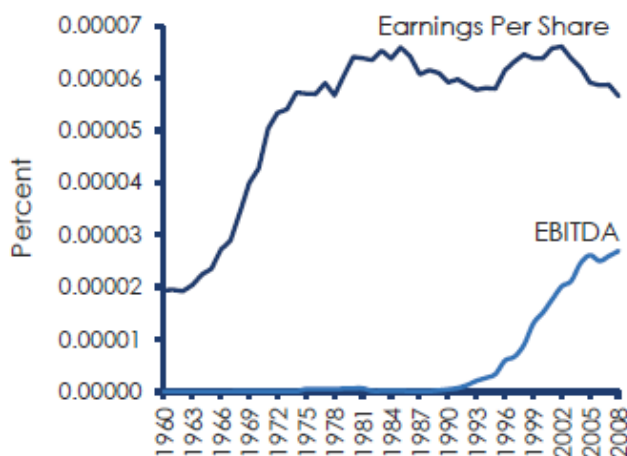
Figura 5: Múltiplo preferido para cada industria

Sector	1st Choice	2nd Choice
Materials	Price/earnings	EV/book value
Industrials	EV/EBITDA	Price/earnings
Telecommunication Services	EV/EBITDA	Price/sales per share
Consumer Discretionary	Price/earnings	PEG
Consumer Staples	EV/EBITDA	Price/earnings
Energy	Price/earnings	Price/cash earnings
Financial	Price/book	Price/earnings
Health Care	Price/earnings	EV/EBITDA
Technology	EV/EBITDA	Price/sales per share
Utilities	Price/earnings	Price/cash earnings

Fuente: Block (2010)

Junto con el PER, el múltiplo de valoración por excelencia es el EV/EBITDA. Mientras que el múltiplo precio-beneficio (P/E) había sido tradicionalmente utilizado en la industria financiera, el múltiplo EV/EBITDA surgió posteriormente. Will Thorndike en su libro *The Outsiders* (2012), atribuye a John Malone, CEO de TCI durante 24 años, la introducción del término EBITDA, cuyo uso se ha extendido y generalizado a pesar de no ser un término definido por la normativa contable US GAAP. La figura 6 muestra la popularidad del beneficio por acción (EPS) y el EBITDA en los libros que Google escanea. Las referencias al EPS se mantienen estables durante los últimos 40 años, mientras que las del EBITDA muestran un alto crecimiento desde que irrumpió en escena a finales de la década de 1980.

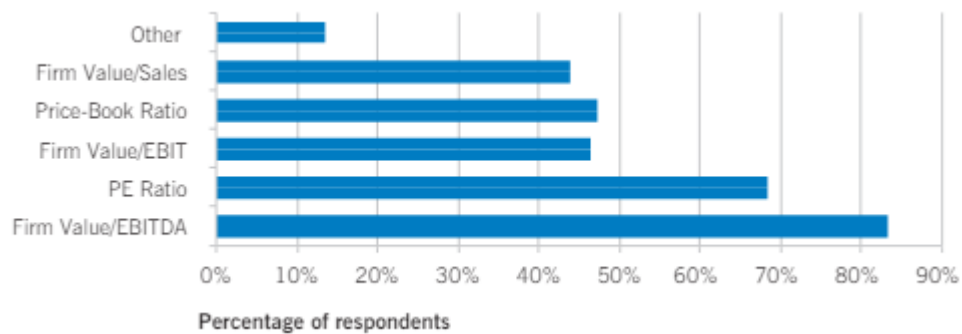
Figura 6: Referencia a “EBITDA” y a “EPS” en libros, 1960-2008



Fuente: Mauboussin (2018)

Así pues, aunque su definición conceptual ocurrió con posterioridad a la del PER, los múltiplos de valor de negocio (EV), y concretamente su relación con el EBITDA se empezaron a utilizar de forma creciente y masiva por inversores y analistas financieros, como se muestra en la figura 7:

Figura 7: Popularidad de los diferentes múltiplos de valoración



Fuente: Bancel y Mittoo (2014)

El EBITDA representa el beneficio antes de intereses, impuestos y amortizaciones⁴. Es considerado a menudo una métrica más apropiada para medir la capacidad operativa de la compañía, eliminando el impacto de la estructura financiera, y también una mejor aproximación a la generación de caja, al no incluir las dotaciones anuales a la amortización de los diferentes activos (aunque eso no significa que el EBITDA sea en sí un flujo de caja, ya que excluye elementos como el CAPEX o la inversión en fondo de maniobra).

Block (2010) recuerda que el EBITDA “se puso de moda” durante el auge del mercado de acciones de internet a finales de la década de 1990. Puesto que la mayoría de las compañías tecnológicas que salían al mercado en aquellos años no tenían beneficios, necesitaban una medida alternativa aparte del Beneficio Neto. Este autor apunta que esto ha conllevado a ciertos críticos a referirse al EBITDA como un “*beautification of earnings*” o incluso como “*financial headfake*”. Block (2010) continúa recordando como Warren Buffet, entre otros, considera el uso del EBITDA “fraudulento” mientras otros lo llaman “*earnings before the bad stuff*”⁵. Puesto que el EBITDA no es una magnitud definida por la normativa contable US GAAP, la forma de calcularlo puede no ser homogénea, y muchas compañías ajustan el EBITDA intentando que este beneficio sea más comparable con el de otras empresas de su

⁴ *Depreciation* se refiere a la amortización de activos físicos o tangibles (*Property, Plant & Equipment*), mientras que *Amortization* hace referencia a la amortización de activos intangibles (patentes, marcas o fondo de comercio)

⁵ Zweig (2018) recuerda que Charles Munger se refería al EBITDA como “*bullshit earnings*.”

sector, o para eliminar el efecto de elementos extraordinarios. Como recuerda Zweig (2018) con cierto sentido del humor : Las variaciones incluyen Ebitdac (con un cambio en los costos de adquisición utilizados por las aseguradoras); Ebitdao (con el coste de las opciones utilizadas para remunerar a los managers); Ebitdap (pensiones y otros beneficios de jubilación); Ebitdar (los costes de alquilar bienes inmuebles o aviones, dependiendo de la industria); Ebitdare (pérdidas, ganancias y otros ajustes en bienes inmuebles); Ebitdas o Ebitdasc (stock options para los gestores)); y Ebitdax (costes de exploración para las compañías de petróleo y gas). ¿Podrán Ebitdada o Ebitdaft estar muy lejos?.

Otra limitación del EBITDA, recogida no solo por Block (2010) sino también por ejemplo por Koller et al. (2015), es que la eliminación de la amortización, a pesar de ser un elemento que no representa movimiento de caja, implica no tener en cuenta las muy necesarias inversiones en Activos de Capital (*capex*), cuya financiación, si uno se fija únicamente en el EBITDA, parece “caída del cielo”. En palabras de Koller et al. (2015), “la amortización de los activos existentes es el equivalente contable a ir preparando un fondo para futuras inversiones de capital que serán necesarias para reemplazar tales activos”⁶. También Chan y Lui (2011) creen que el D&A (*depreciation y amortisation*) reflejan el *capex* de la empresa en años anteriores y orientan a inversores sobre el crecimiento en beneficios y la sostenibilidad de futuras inversiones, por lo que defienden que el EV/EBIT es un múltiplo mejor que el EV/EBITDA.

A pesar de estos inconvenientes, no puede negarse la enorme y creciente popularidad del EBITDA, que no se limita a analistas financieros e inversores, sino también en ámbitos de *reporting* financiero contable. Bouwens et al. (2019) analizaron decenas de miles de informes anuales y comunicados de resultados de empresas del S&P 1500 entre 2005 y 2016 y comprobaron que alrededor del 25% de las mismas hacen referencia al EBITDA en su informe anual, y que un 18% de las noticias de prensa también lo mencionan, a pesar que, como ya se ha mencionado, es una magnitud no incluida en los principios contables (GAAP). Según este estudio, las empresas que enfatizan el EBITDA son, en promedio, más pequeñas, más apalancadas, más intensivas en capital y menos rentables que sus pares. Según Rozenbaum (2019) el uso del EBITDA en los mercados de capitales y la divulgación de

⁶ Estos autores dejan, por tanto, la depreciación de activos tangibles, pero eliminan la de activos intangibles, ya que creen que el reemplazo de los activos intangibles sí está siendo ya considerado en gastos como marketing y gastos comerciales. En su opinión, el múltiplo más correcto sería EV/EBITA (antes de *amortization*, pero después de *depreciation*), prefiriéndolo al EV/EBITDA o al EV/EBIT.

EBITDA en los anuncios de resultados es prevalente, habiendo aumentado las empresas que publican dicha cifra del 17% en 2003 al 35% en 2011.

Mauboussin (2018) recoge distintas razones de la popularidad del EBITDA frente a métricas más tradicionales: es una aproximación al flujo de caja operativo y por tanto indica la capacidad para invertir y pagar deudas; puede ser útil para analizar empresas que pierden dinero y tienen beneficio negativo pero EBITDA positivo; y el EBITDA parece ser más adecuado para comparar empresas con diferentes niveles de deuda, ya que no incluye los gastos financieros.

También Koller et al (2015) consideran los múltiplos de EV superiores al PER, aduciendo que eliminan las distorsiones causadas por distintas estructuras financieras y por activos no operativos que generen ingresos o gastos no operativos recogidos en el Beneficio Neto. Sin embargo, como ya se ha mencionado, refinan el enfoque del EV/EBITDA, inclinándose más por el EV/EBITA (es decir, excluyendo *depreciation*, la dotación anual a la amortización de activos tangibles, pero dejando *amortization*, la dotación anual a la amortización de activos intangibles) o el EV/NOPAT (*Net Operating Profit After Taxes*, o beneficio operativo después de impuestos). Chan y Lui (2010) afirman que el EV/EBITDA produce resultados menos sesgados en industrias en las que los flujos de caja son relevantes pero las inversiones en capex son altas e irregulares, como telecomunicaciones, energía o industria.

En parte por estas razones, el EV/EBITDA se convirtió en el principal múltiplo de valoración que los inversores de *Private Equity* utilizaban para evaluar LBOs en la década de 1980 y sigue siendo el enfoque más utilizado en las operaciones de Capital Riesgo. A medida que las inversiones “alternativas” en *Hedge Funds* o en Capital Riesgo han ido aumentando, también lo ha hecho la aplicación del EV/EBITDA. El EV/EBITDA aparecen también muy frecuentemente en los modelos DCF. La mayoría de estos tienen un período de estimación explícita de flujos de caja libres, y un valor residual o terminal, para reflejar los flujos más allá de dicho período explícito. Según Gompers et al (2015), las compañías de *Private Equity* modelan proyecciones explícitas de flujos de caja durante cinco años y la mayoría utilizan múltiplos de EBITDA para estimar el valor terminal. Nissim (2019) confirma que es el múltiplo de EBITDA el más usado en los cálculos del valor terminal en un DCF, por encima de otros que él encuentra más sólidos conceptualmente (como el NOPAT), e indica que este valor terminal puede representar el 70-80% del valor empresarial. Vidal-García y Ribal (2019) comprueban empíricamente la validez del múltiplo EV/EBITDA para estimar el valor terminal de un DCF en una muestra de compañías no cotizadas de la industria

agroalimentaria española. En cierto modo, los flujos de caja estimados explícitamente resultan no ser más que el “preludio” de la decisión principal: qué múltiplo aplicar al EBITDA del último año⁷.

2.3 Algunos estudios de valoración por múltiplos

Además de su uso cotidiano en la práctica diaria de los mercados, los múltiplos de valoración son objeto de numerosos estudios académicos, algunos ejemplos de los cuales se mencionan a continuación.

Una ventaja del EBITDA sobre el Beneficio Neto que se suele esgrimir es que el Beneficio Neto es más manipulable y más dependiente de políticas contables que el EBITDA. En Zheng y Stangeland (2007) se estudia el precio y evolución de salidas a bolsa en relación al crecimiento anunciado en ventas, EBITDA y beneficios, y concluyen que la calidad de la compañía está relacionada con las ventas y con el EBITDA, y no tanto con los beneficios, ya que las dos primeras magnitudes son más difíciles de manipular que la última, más sujeta a lo que se conoce como “*earnings management*”.

Liu et al (2001) analizan una variedad de múltiplos de valoración, tanto de *Equity* como de *Enterprise Value*, tratando de ver cuáles explican de mejor manera los precios de las acciones. Utilizan variables históricas (flujo de caja operativo, EBITDA, ventas, beneficio y valor teórico contable del *Equity*) así como variables estimadas, basadas en proyecciones de analistas (beneficio por acción de los próximos dos años, o crecimiento del beneficio) y construyen los múltiplos por industria y también para el universo total de empresas consideradas. Deng et al. (2010) usan también ambos tipos de múltiplos, de equity y de EV contruidos con datos fundamentales de las empresas, pero en lugar de aplicarlos a empresas establecidas y rentables, amplían la muestra incluyendo empresas con pérdidas, *startups* o empresas con pérdidas. Payne et al (2018) aplican ratios de EV/EBITDA a empresas en mercados emergentes, y Loughran y Wellman (2011) concluyen que los múltiplos de valoración de EV son determinantes para explicar los retornos en los precios de las acciones.

En las siguientes secciones se volverá a hacer referencia a estos trabajos y a cómo definen cada uno de los múltiplos que utilizan, así como su relación con la base de datos compustat que se utiliza en este TFG, un resumen de todo esto se recoge en la tabla 1 a continuación.

⁷Pueden encontrarse intentos conceptualmente válidos para lidiar con los problemas del valor terminal, como Holland (2018), aunque la mayoría de los inversores no los utilizan en la práctica diaria.

Tabla 1. Artículos académicos sobre múltiplos de valoración: definición de las variables principales

Trabajo	Bases de datos utilizadas	Clasificación industrias	Variables utilizadas	Definición de variables
Liu, Nissim & Thomas (2002)	Compustat CRSP IBES	IBES	EV per share	Book Value of Debt = l/t debt (#9) + debt in current liabilities (#34) + preferred stock (#130) - preferred treasury stock (#227) + preferred dividends in arrears (#242) TP (EV per share) = Book Value of Debt / shares outstanding (#25) + share price (P)
			BV per share	Book equity (#60) / shares outstanding (#25)
			Sales per share	Sales (#12) / shares outstanding (#25)
			EBITDA per share	EBITDA (#13) / shares outstanding (#25)
			Actual EPS Compustat	EPS excluding extraordinary items (#58)
			Actual EPS IBES	
			EPS1	mean of IBES one year EPS forecast
			EPS2	mean of IBES two year EPS forecast
Payne, Wong & Tyler (2018)	Damodaran 2014		EV	EV = Market Capitalization + Debt + Preferred Stock + Minority Interest - Cash and equivalents
			EBITDA	Earnings Before Interest, Taxes, Depreciation and Amortization
Loughran & Wellman (2011)	Compustat CRSP	SIC	Enterprise Multiple (EM) = EV / EBITDA	EV = market value of equity + prior year's total debt (DLC and DLTT) + preferred stock (PSTKRIV) - cash and s/t investments (CHE) EBITDA = prior year's operating income before depreciation (OIBDP)
			Size: Market capitalization / Book to Market	Stock Price * shares outstanding
			Earnings to Price	[Prior's year equity income (income before extraordinary, IB) - dividend on preferred stock (DVP) + income statement deferred taxes (TXDI)] / market value of equity as of Dec of previous year.
Platt, Demirkan & Platt (2010)	Compustat CRSP	SIC (#324)	EV (last trading day of the year)	EV = MarketCap (share price #199*nº shares #25) + Debt (l/t #9 + s/t #44 + s/t #206) - Cash (#1)
			Estimated EV (EEV)	EEV = PV of CCF + TV CCF = net income (#172)+depreciation (#14)-CAPEX (#128)+WC-LWC+deferred taxes (#74)-Ldata74+interest paid (#15) - interest received (#62) WC = net current assets (#4-#5) - cash (#1) + notes (#206) LWC is lagged value of WC and Ldata74 is lagged value for #74
Zheng & Stangeland (2007)	Compustat CRSP Securities Data corporation (SDC)		growth in net sales	net sales (#12)
			growth in EBITDA	EBITDA (#13)
			Income before extraordinary Accruals	Earnings (#237)
Lifland (2011)	Compustat	GICS	EV	EV = Market Capitalization + Debt + +Non-controlling minority interest+Preferred Shares - Excess Cash
Deng, Easton & Yeo (2010)	Compustat CRSP	SIC	EV	EV = MV + net financial obligations
			Market Value	MV = closing price for the fiscal year * number of shares outstanding (compustat #25)
			Net financial obligations	Financial Liabilities - Financial Assets
			Financial Liabilities	debt in current liabilities (#34)+ l/t debt (#9) + preferred stock (#130) + preferred dividends in arrears (#242) + minority interest (#38) - preferred treasury stock (#227)
			Financial Assets	Cash and s/t investments (#1) + investments and advances (#32)
			EBITDA	Earnings Before Interest, Taxes, Depreciation and Amortization (#13)
			Net Income before extrao	NIBE = net income (#172) - extraordinary items and discontinued operations (#48)

Fuente: Elaboración propia

Un trabajo clave a la hora de entender las discrepancias que van a intentar identificarse en eses TFG es el de Jullens (2012), que se encuadra en las famosas publicaciones de UBS: *The Footnotes Compendium*. Estos *footnotes* son pequeños trabajos de análisis realizados por profesionales de la industria financiera en el que plantean algunos de los problemas cotidianos que se encuentran en la práctica diaria de los mercados. En el caso del referido *footnote* (Jullens, 2012), trata de demostrar la relación algebraica que hay entre el EV/EBITDA y el PER, es decir, los ajustes que es necesario aplicar sobre un múltiplo para que se genere el otro. Esta forma de reconciliar ambos múltiplos pone de manifiesto las variables de las que dependen los múltiplos y que pueden en un momento determinados hacer que uno esté por encima de sus comparables y otro por debajo. Es decir, puede hacer que la acción esté “...expensive on a PE basis but cheaper on EV/EBITDA” (o a la inversa).

Una vez identificados los ajustes y las variables financiero-contables de cada uno, se calcularían los “factores de ajuste” que sumados al EV/EBITDA nos proporcionarían el valor del PER. En su trabajo, Jullens (2012) realiza este trabajo de reconciliación con los múltiplos de la industria de destilerías y bebidas alcohólicas europeas.

Tabla 2. Reconciliación del múltiplo EV/EBITDA y el PER

EV/EBITDA
Depreciation adjustment factor
Amortisation adjustment factor
Non-core asset adjustment factor
Gearing adjustment factor
Taxation adjustment factor
<u>Minority interest adjustment factor</u>
PER

Fuente: Jullens (2012)

Estos ajustes son fundamentales a la hora de entender y explicar las discrepancias en los ratios de valoración: diferentes estructuras de apalancamiento, diferentes amortizaciones de activos (so sólo debido a políticas contables sino a distintos planes de inversión), y diferencias en los impuestos, inversiones en activos no afectos o minoritarios, producirían múltiplos desalineados entre una empresa y lo que se habrían considerado sus comparables.

CAPÍTULO III: METODOLOGÍA: CONSTRUCCIÓN DE LA MUESTRA Y CÁLCULO DE LOS MÚLTIPLOS

3.1 Bases de datos utilizadas

Este TFG va a utilizar principalmente datos de Compustat, tanto para la información contable financiera, como para los precios de cotización.

Compustat es una base de datos de información financiera y de mercado sobre empresas, índices e industrias globales. El tipo de información publicada por Compustat incluye Estándares de Clasificación Industrial Global (GICS), datos de precios, datos de ganancias, participaciones internas e institucionales, y otra información dirigida a inversores, analistas, economistas, académicos y otros investigadores. La base de datos está publicada por Standard and Poor's desde 1962, con información que se remonta a 1950.

El servicio, que se nutre principalmente de información presentada ante la SEC⁸, compila informes financieros presentados por empresas cotizadas, tanto anuales como trimestrales, incluyendo ingresos, gastos, activos y pasivos e informes históricos de los mismos datos. El historial de precios de las acciones también está disponible.

Compustat es uno de los proveedores de información líder para profesionales, inversores y académicos del mercado financiero, como demuestra el hecho de que la mayor parte de los trabajos académicos consultados elaboran los múltiplos cruzando información de dos bases de datos diferentes: Compustat para los datos contables y de fundamentales de las empresas, y CRSP para los datos diarios de precios de cotización de las acciones (ver por ejemplo Deng et al. (2010); Lifland (2011); Liu et al. (2002); Loughran y Wellman (2011); Platt et al. (2010) o Zheng y Stangeland (2007)).

3.2 Clasificación por industria: códigos SIC

La forma en que se van a agrupar las empresas de compustat en diferentes sectores o industrias es una cuestión fundamental puesto que van a constituir el benchmark o universo de comparables respecto al cual se a realizar la valoración y a determinar si teóricamente el precio de la acción estaría sobrevalorado o infravalorado. Este TFG va a agrupar las empresas utilizando los códigos SIC, concretamente el código SIC de dos dígitos.

⁸ *Securities and Exchange Comission*, autoridad norteamericana de supervisión de los mercados, equivalente a la CNMV española.

Los códigos SIC (de las siglas en inglés: *Standard Industrial Classification*) son un sistema en el que se clasifican industrias con un código de cuatro dígitos para estandarizar y ordenar los distintos sectores industriales con fines estadísticos. Establecido en Estados Unidos en 1937, lo utilizan las agencias gubernamentales norteamericanas para clasificar áreas industriales. El sistema SIC fue revisado por última vez en 1987. La Oficina del Censo lo utilizó por última vez para el Censo Económico de 1992, y ha sido reemplazado por el Sistema de Clasificación Industrial de América del Norte (código NAICS) publicado en 1997. Sin embargo, los códigos SIC continúan utilizándose en varios departamentos y agencias del gobierno norteamericano, entre ellos la autoridad supervisora de los mercados financieros (SEC), así como en el mundo empresarial y académico. En la página web de la SEC, se explica que

“Los códigos SIC que aparecen en los archivos EDGAR⁹ depositados por una empresa, indican el tipo de negocio de dicha empresa. Estos códigos también se utilizan en la División de Finanzas Corporativas como base para asignar la responsabilidad de supervisión de la documentación depositada por la empresa. Por ejemplo, el personal de la Oficina de Energía y Transporte revisaría la documentación de una empresa cuyo negocio fuera Minería Metálica (SIC 1000).”¹⁰

El sistema SIC clasifica la economía en 11 grandes divisiones que pueden consultarse en la tabla 3:

⁹ EDGAR, de las siglas en inglés de *Electronic Data Gathering, Analysis, and Retrieval*, es el principal sistema para el depósito y registro de documentación por parte de empresas y otros agentes que depositan esta información para cumplir con los requisitos legales de distintas normas aplicables. EDGAR gestiona millones de presentaciones de empresas e individuos, beneficiando a inversores, empresas y a la economía estadounidense en su conjunto al aumentar la eficiencia y transparencia de los mercados de valores. El sistema procesa alrededor de 3.000 presentaciones por día, ofrece 3.000 terabytes de datos al público anualmente y admite 40.000 nuevos contribuyentes por año en promedio.

¹⁰ <https://www.sec.gov/corpfin/division-of-corporation-finance-standard-industrial-classification-sic-code-list>.

Tabla 3. Divisiones industriales según el SIC de un dígito

Agricultura, silvicultura y pesca
• Minas
• Construcción
• Industria
• Transporte y servicios públicos
• Comercio al por mayor
• Comercio al por menor
• Finanzas, seguros, bienes raíces
• Servicios
• Administración pública
• Establecimientos no clasificables

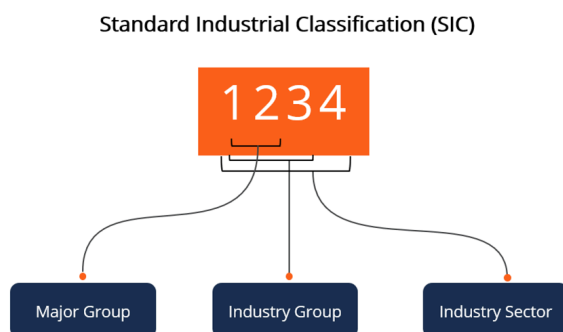
Fuente: elaboración propia a partir de:

<https://www.sec.gov/corpfin/division-of-corporation-finance-standard-industrial-classification-sic-code-list>

Estos se dividen a su vez en 83 grupos principales de dos dígitos y luego se subdividen en 416 grupos industriales de tres dígitos y en más de 1.000 industrias de cuatro dígitos.

Los códigos SIC tienen una estructura jerárquica de descendente, comenzando con las características generales y continuando hacia las más específicas. Cada empresa tiene un código SIC principal que indica su principal línea de negocio, una de las 11 divisiones generales a las que se ha hecho referencia. Así, los dos primeros dígitos del código representan el principal sector industrial al que pertenece la empresa, y el tercer y cuarto dígito describen la subclasificación más detallada del grupo empresarial y la especialización, respectivamente.

Figura 8.: Composición de los dígitos del código SIC



Fuente: <https://www.anfagua.es/clasificacion-industrial-estandar-sic/>

Por ejemplo, "36" se refiere a una empresa que se dedica a "equipos electrónicos y otros", el "7" como tercer dígito para obtener "367" indica que la empresa opera en "Electrónica, Componentes y Accesorios", y el cuarto dígito distingue la especialización industrial más específica, por lo que un código "3672" concretaría que la empresa se dedica a "Placas de circuito impreso".

Como se ha mencionado, este TFG va a ordenar las empresas por industrias según los dos primeros dígitos del código SIC, los que indican el principal sector industrial de la empresa, y que pueden consultarse en la tabla 4:

Tabla 4: Códigos SIC-dos dígitos

01 Agricultural Production - Crops	34 Fabricated Metal Products Manufacturing	60 Depository Institutions
02 Agricultural Production - Livestock	35 Industrial & Commercial Machinery Manufacturing	61 Non-Depository Credit Institutions
07 Agricultural Services	36 Electronic & Other Electrical Equipment Manufacturing	62 Security & Commodity Brokers
08 Forestry	37 Transportation Equipment Manufacturers	63 Insurance Carriers
09 Fishing Hunting & Trapping	38 Measuring & Analyzing Instruments Manufacturers	64 Insurance Agents Brokers & Services
10 Metal Mining	39 Miscellaneous Manufacturing Industry	65 Real Estate
12 Coal Mining	40 Railroad Transportation	67 Holding & Other Investment Offices
13 Oil & Gas Extraction	41 Local/Suburban Transit & Hwy Passenger	70 Hotels Rooming Houses & Camps
14 Mining & Quarrying - Non-metallic Minerals	42 Motor Freight Transportation Warehouse	72 Personal Services 73 Business Services
15 Building Construction - General Contractors	43 Postal Service	75 Auto Repair Services & Parking
16 Heavy Construction - except Building	44 Water Transportation	76 Miscellaneous Repair Service
17 Construction - Special Trade Contractors	45 Transportation by Air	78 Motion Pictures
20 Food & Kindred Products Manufacturing	46 Pipelines - except Natural Gas	79 Amusement & Recreation Services
21 Tobacco Products Manufacturing	47 Transportation Services	80 Health Services
22 Textile Mill Products Manufacturing	48 Communications	81 Legal Services
23 Apparel & Other Finished Products - Manufacturing	49 Electric Gas & Sanitary Services	82 Educational Services
24 Lumber & Wood Prods - except Furniture Manufacturing	50 Wholesale Trade - Durable Goods	83 Social Services
25 Furniture & Fixtures Manufacturing	51 Wholesale Trade - Nondurable Goods	84 Museums Art Galleries & Gardens
26 Paper & Allied Products Manufacturing	52 Building Materials & Hardware	86 Membership Organizations
27 Printing Publishing & Allied Industries	53 General Merchandise Stores	87 Engineering & Accounting & Management Services
28 Chemicals & Allied Prods Manufacturing	54 Food Stores	89 Miscellaneous Services NEC
29 Petroleum Refining & Related Industry Manufacturing	55 Automotive Dealers & Service Stations	91 Executive Legislative & General Government
30 Rubber & Miscellaneous Plastics Manufacturing	56 Apparel & Accessory Stores	92 Justice Public Order & Safety
31 Leather & Leather Products Manufacturing	57 Home Furniture & Furnishings Stores	95 Admin-Environmental Quality Programs
32 Stone Clay Glass & Concrete Products Manufacturing	58 Eating & Drinking Places	96 Administration of Economic Programs
33 Primary Metal Industries Manufacturing	59 Miscellaneous Retail	97 National Security & International Affairs
		99 Non-classified establishments

Fuente: elaboración propia a partir de:

<https://www.sec.gov/corpfin/division-of-corporation-finance-standard-industrial-classification-sic-code-list>

En la tabla anterior se observa que algunos de los dígitos faltan, por ejemplo se pasa del código 02 al código 97. Estas omisiones reflejan cambios en la industria y la naturaleza evolutiva de las actividades económicas, debiéndose principalmente a los siguientes motivos:

Exclusión intencional: Cuando se desarrolló el sistema SIC, ciertos códigos de dos dígitos se dejaron sin asignar intencionalmente con el fin de permitir futuras expansiones o porque no se identificaron grupos industriales adecuados en ese momento.

Redundancia y consolidación: Con el tiempo, algunas industrias se consolidaron o reclasificaron, lo que resultó en la eliminación de ciertos códigos de dos dígitos. Por ejemplo,

si una industria se redujo o se fusionó con otra, su código específico podría haber sido eliminado o absorbido por otra categoría.

Evolución de la industria: El sistema SIC ha pasado por varias revisiones desde su implementación inicial. Durante estas revisiones, se eliminaron industrias que ya no existían o que se consideraron irrelevantes, lo que llevó a vacíos en la secuencia.

3.3 Definición de múltiplos y variables compustat utilizadas

a. Variables históricas TTM (*Trailing Twelve Months*)

En este TFG, se van a construir los múltiplos utilizando precios diarios y variables contables históricas ya reportadas por las empresas. Si bien para ejercicios de valoración es frecuente utilizar proyecciones futuras del beneficio neto o del EBITDA, son también muchos los trabajos académicos y análisis de mercado que utilizan información histórica, entre ellos Loughran y Wellman (2011); Platt et al. (2010) o Deng et al. (2010).

Concretamente, se van a utilizar variables conocidas como TTM (por sus siglas en inglés *Trailing Twelve Months*). Este término describe los últimos 12 meses consecutivos de los datos financieros o contables de una empresa. El cálculo de las cifras TTM es posible puesto que muchas empresas, especialmente las cotizadas, notifican de sus resultados a la autoridad de mercado competente con una mayor regularidad que simplemente una vez al año, reportando a menudo resultados trimestrales. De esta forma, el último beneficio TTM de una empresa se obtendría mediante la suma de los últimos cuatro beneficios trimestrales reportados. Así, los 12 meses estudiados no van a coincidir necesariamente con el cierre del año, únicamente cuando los últimos cuatro beneficios reportados sean todos del mismo ejercicio económico.

Las cifras TTM representan más adecuadamente la situación actual de la empresa que por ejemplo los informes anuales presentados a cierre del ejercicio económico, y que según en que momento se esté haciendo el análisis pueden suponer información de hace meses o casi un año. El cálculo de las cifras de TTM se realiza para varias métricas de la cuenta de resultados, incluidas ventas, EBITDA o beneficios. Es frecuente utilizar los beneficios TTM para calcular ratios financieros. Por ejemplo la relación precio/beneficio (múltiplo PER), se puede calcular como el precio actual de la acción dividido por las ganancias por acción (Beneficio por Acción o BPA) de los últimos 12 meses de la empresa, o como la capitalización bursátil actual entre el beneficio neto de los últimos 12 meses.

Con el fin de calcular las cifras TTM para la construcción de los múltiplos, se ha buscado en la base de datos compustat, no las variables anuales, sino su equivalente en variables trimestrales.

b. Definición de múltiplos y variables de compustat

Si, como ya se ha explicado en secciones anteriores, el PER se calcula como el *equity* (a valor de mercado) entre el Beneficio Neto, y el otro múltiplo de valoración de este estudio se calcula como el EV dividido por el EBITDA, vamos a proceder a definir cada uno de los componentes de los múltiplos, mencionando cómo se han definido en algunos trabajos consultados, y detallando qué variables de compustat se utilizarán en el TFG.

Equity (capitalización bursátil o market cap): Numerador del ratio PER

El numerador del ratio PER es el valor de mercado de las acciones de la empresa, de sus recursos propios, y representa la riqueza de los accionistas, que son en última instancia los dueños de la compañía. En el caso de empresas cotizadas, y hay que recordar que todas las consideradas en este TFG son cotizadas, el valor de mercado del *equity* es el producto del precio de cotización de la acción por el número de acciones en circulación, lo que se conoce como capitalización bursátil (*Market Cap*).

Aunque hay cierto debate en ocasiones sobre si el número de acciones debe ser el *fully diluted*¹¹, o sobre si usar el número medio de acciones durante el periodo que se considere, la mayor parte de los trabajos consultados (ver por ejemplo Liu et al., 2002; Loughran & Wellman, 2011; Platt et al., 2010; Lifland, 2011 o Deng et al., 2010) utilizan el número de acciones reportado por la propia compañía en sus estados financieros anuales o a cierre del periodo en cuestión (trimestre).

En cuanto al precio, la práctica habitual es utilizar el precio de la acción a cierre de mercado en una fecha determinada. Aunque el trabajo de Platt et al (2010) utiliza únicamente el precio de cierre del último día del año fiscal para calcular la capitalización bursátil, es mucho más común utilizar precios diarios e ir por tanto calculando valores diarios del *equity* para elaborar los múltiplos de valoración.

En este TFG se va a calcular el *equity* (numerador del ratio PER) multiplicando el precio diario de la acción a cierre de mercado por el número de acciones en circulación reportado por

¹¹ “Fully diluted number of shares” (número de acciones totalmente diluido) se refiere al total de acciones en circulación de una empresa si todas las opciones sobre acciones, *warrants*, bonos convertibles y otros valores convertibles en acciones ordinarias fueran ejercidos o convertidos.

la empresa a cierre del periodo. Las variables de compustat utilizadas (con su clave) son *closing market price* (prccd)¹² y *common shares outstanding* (cshoq):

*Equity (Market Cap) = closing market price (prccd) * number of shares outstanding (cshoq)*

Enterprise Value: Numerador del ratio EV/EBITDA

Desde el punto de vista de los activos, el *Enterprise Value* (EV) es el valor de mercado de la actividad principal de la empresa, o el valor de los activos en cuanto que generan esa actividad. Quedan excluidos activos no operativos, como el exceso de caja o las filiales consolidadas por puesta en equivalencia. Desde el punto de vista del capital invertido, el EV es la suma de la capitalización bursátil (valor de mercado del *equity*), la deuda financiera a largo y a corto plazo, equivalentes de deuda, acciones preferentes e intereses minoritarios, restando posteriormente el exceso de caja y los activos no operativos.

Muchos trabajos académicos utilizan múltiplos de EV, siendo más o menos detallado en la definición del mismo. La definición de EV como suma de la capitalización bursátil y la deuda neta, aunque utilizada a veces en la práctica, o en trabajos como Chan & Lui (2010) o Platt et al (2010), es una simplificación que debe ser tratada con cautela. Más detallada es la definición de EV de trabajos como Loughran and Wellman (2001) o Payne et al. (2018), en los que el EV es la suma de la capitalización bursátil más la deuda, las acciones preferentes y los intereses minoritarios, y menos la caja y los equivalentes de caja. Liffland (2011) sigue el mismo enfoque, e insiste en que la caja y equivalentes que se deben restar al cálculo son el *excess cash* que sobraría tras hacer frente a los pasivos corrientes.

Dang et al (2010) expanden aún más el cálculo del EV, añadiendo a los elementos ya mencionados los dividendos preferentes atrasados (sumando) y las acciones preferentes en autocartera (restando). Liu et al. (2001) calculan el EV por acción, dividiendo el valor contable de la deuda (incluyendo largo plazo, corto plazo, acciones preferentes y dividendos preferentes atrasados, y excluyendo autocartera) entre el número de acciones y sumándole después el precio por acción.

En resumen, podemos afirmar que de los trabajos consultados se extrae la siguiente definición de EV, con las variables de compustat (y sus códigos numéricos)¹³ en términos anuales:

¹² Recuérdese que la mayoría de los trabajos consultados utilizan CRSP como fuente de información de precios, y no Compustat como el presente TFG.

¹³ Tal y como aparecen en los artículos consultados.

$EV = Equity (Market Cap) + Financial Liabilities - Financial Assets$

$Equity (Market Cap) = closing\ market\ price\ (prccd) * number\ of\ shares\ outstanding\ (cshoq)$

$Financial\ Liabilities = debt\ in\ current\ liabilities\ (\#34) + l/t\ debt\ (\#9) + preferred\ stock\ (\#130) + preferred\ dividends\ in\ arrears\ (\#242) + minority\ interest\ (\#38) - preferred\ treasury\ stock\ (\#227)$

$Financial\ Assets = Cash\ and\ s/t\ investments\ (\#1) + investments\ and\ advances\ (\#32)$

Puesto que este TFG va a utilizar variable TTM, tal y como se ha explicado anteriormente, vamos a utilizar las variables trimestrales reportadas por las compañías, y no las anuales. Estas variables trimestrales pueden también obtenerse de compustat, aunque las empresas no reportan trimestralmente todas las magnitudes que reportan anualmente.

Trasladando el cálculo del EV extraído de los trabajos académicos fórmulas equivalentes con variables trimestrales anualizadas a cifras TTM, es necesario prescindir de algunos elementos que no están disponibles en cifras trimestrales, concretamente *preferred dividends in arrears* (#242), *preferred treasury stock* (#227) e *investments and advances* (#32). Creemos que son variables de cuantías residuales o nulas la mayoría de las veces, y que la ventaja de utilizar cifras TTM compensa las desventajas de prescindir de ellas¹⁴.

Por tanto, en este TFG va a calcular el *Enterprise Value* (numerador del ratio EV/EBITDA) como el valor de mercado del *equity* más la deuda a largo plazo, a corto plazo, las acciones preferentes y los intereses minoritarios y menos la caja y equivalentes. Las variables de compustat utilizadas (con su clave) son:

$Equity (Market Cap) = closing\ market\ price\ (prccd) * number\ of\ shares\ outstanding\ (cshoq)$

$EV\ adjustment = debt\ in\ current\ liabilities\ (dlcq) + l/t\ debt\ (dlttq) + preferred\ stock\ (pstq) + minority\ interest\ (mibq) - Cash\ and\ s/t\ investments\ (cheq)$

$EV = Market\ Cap + EV\ adjustment$

Puesto que las partidas del cálculo del EV son magnitudes del Balance, la cifra TTM no resulta de sumar las de los cuatro últimos trimestres, sino que se utiliza la del último trimestre reportado por la empresa, entendiendo que es la que mejor refleja la situación actual de dicha magnitud.

¹⁴ Nótese que con las cifras de Balance, como son las del cálculo del EV, la cifra TTM no resulta de sumar las de los cuatro últimos trimestres, sino que se utiliza la del último trimestre reportado por la empresa, entendiendo que es la que mejor refleja la situación actual de dicha magnitud.

EBITDA: Denominador del ratio (EV/EBITDA)

EBITDA es un acrónimo de la expresión en inglés *Earnings Before Interest, Taxes, Depreciation and Amortization*. En el capítulo anterior se ha hecho referencia a algunos de los motivos por el que esta magnitud ha ido haciéndose cada vez más popular: permite comparar empresas de negocios similares pero estructuras financieras diferentes; al contrario que la cifra de ventas, incluye la estructura de costes y los márgenes operativos; y se centra en la generación de caja al no incluir las dotaciones anuales a la amortización. Sin embargo, también se mencionaban algunos de los problemas de utilizar el EBITDA, como el hecho de que excluyen costes o inversiones importantes y pueden dar lugar a una visión edulcorada de los beneficios, así como la falta de consistencia en su cálculo, al no ser una magnitud recogida en los principios contables (GAAP).

En este TFG se va a tomar EBITDA (denominador del ratio EV/EBITDA) como beneficio operativo (que es previo a intereses e impuestos) antes de las dotaciones anuales a amortizaciones de activos. La variable trimestral de compustat utilizada (con su clave) es:

EBITDA = operating income before depreciation (oibdpq)

Se tomará la cifra TTM sumando con la función *rolling* la variable reportada en los últimos cuatro trimestres.

Earnings (beneficio neto): denominador del ratio PER

Zheng y Stangeland (2007) utilizan las ganancias antes de resultados extraordinarios disponibles para los accionistas ordinarios (variable anual de compustat con su código: *Earnings* (#237)). Deng et al. (2010) utilizan el mismo concepto de beneficio neto antes de elementos extraordinarios, pero lo construyen ellos mismos eliminando el efecto los resultados extraordinarios del beneficio del ejercicio (variables anuales de compustat con sus códigos: *net income* (#172) - *extraordinary items and discontinued operations* (#48)) . Loughran y Wellman (2011) lo refinan aún más, ajustando el beneficio antes de extraordinarios con el dividendo de las acciones preferentes y los impuestos diferidos (variables anuales de compustat con sus claves: *income before extraordinaries* (IB) - *dividend on preferred stock* (DVP) + *income statement deferred taxes* (TXDI).

En este TFG se va a tomar el beneficio neto (denominador del ratio PER) como el beneficio neto antes de tener en cuenta resultados extraordinarios, tal y como aparece en compustat, siendo esta variable trimestral (con su clave):

Beneficio neto = income before extraordinary items - available for common (ibcomq)

Se tomará la cifra TTM sumando con la función *rolling* la variable reportada en los últimos cuatro trimestres.

La tabla 13 resume las variables de compustat en términos anuales que utilizan los papers consultados, con sus códigos numéricos, los equivalentes trimestrales que utilizaremos en este TFG para calcular los TTM, con su correspondiente código y clave de compustat. Como ya se ha mencionado, aparecen más variables de las que este TFG utilizará en sus cálculos, ya que, por un lado, los estudios académicos no siempre son homogéneos, y por otro lado, solo algunas variables se reportan trimestralmente y serán utilizadas para calcular los TTM: Todas las variables de la tabla 5 han sido utilizadas en alguno de los estudios de la Tabla 1.

Tabla 5: variables anuales y trimestrales de compustat para construir múltiplos

	Anual	Trimestral
Capitalización bursátil		
<i>common shares outstanding</i>	#25	#61 cshoq
<i>closing-market-price</i>		prccd
Deuda y otros liabilities		
<i>debt in current liabilities</i>	#34	#45 dlcq
<i>debt - due in one year</i>	#44	n.a.
<i>notes payable</i>	#206	n.a.
<i>long term debt - total</i>	#9	#51 dlttq
<i>preferred stock - carrying value</i>	#130	#55 pstkq
<i>dividends-preferred-in arrears</i>	#242	n.a.
<i>minority interest (Balance Sheet)</i>	#38	#53 mibq
<i>treasury stock (dollar amount) - preferred</i>	#227	n.a.
Caja		
<i>cash and short term investments</i>	#1	#36 cheq
<i>investment and advances - equity method</i>	#32	n.a.
EBITDA		
<i>operating income before depreciation</i>	#13	#21 oibdpq
Beneficio neto		
<i>net income (loss)</i>	#172	#69 niq
<i>extraordinary items and discontinued operations</i>	#48	#26 xidoq
<i>EPS (Basic) excluding extraordinary items</i>	#58	#19 epsfxg
<i>income before extraordinary items - available for common</i>	#237	#25 ibcomq

Fuente: elaboración propia

3.4. Universo de compañías

Este trabajo analiza la evolución de los múltiplos de valoración de 5.670 empresas norteamericanas cotizadas, de 70 industrias, y de las que se han obtenido una serie de variables financiero contables fundamentales para construir el PER y el EV/EBITDA diario durante un periodo de unos 13 años, desde el 4 de enero de 2010 hasta el 29 de diciembre de 2023. Se explica a continuación la forma en que se llegó a esta muestra en concreto.

Se generó en primer lugar un *dataset*¹⁵ con los datos contables fundamentales reportados trimestralmente mencionados en la sección anterior ('cshoq', 'cheq', 'dlcq', 'dlttq', 'epsfxq', 'ibcomq', 'mibq', 'niq', 'oibdpq', 'pstkq', 'xidoq'), además de su identificador compustat ('gvkey'), el nombre de la empresa ('conm'), la fecha a la que se refieren los datos, el año fiscal y el trimestre ('datadate', 'fyearq', 'fqtr'), la fecha de publicación o reporte de estos resultados ('rdq'), y códigos de clasificación industrial ('cusip', 'gind', 'gsubind', 'naics', 'sic').

Debido a la gran cantidad de datos de compustat, estas variables se obtenían en *chunks* de 100.000 filas para poder leer el archivo de forma manejable. En un primer momento se incluyen 42.104 empresas desde el 31 de enero de 1962 hasta el 31 de enero de 2024.

A continuación se ajusta el formato de las fechas para que sea homogéneo y todas las empresas recojan los mismos trimestres. Se procede entonces a ajustar las observaciones duplicadas, y a eliminar las empresas con observaciones incompletas, es decir, aquellas a las que les faltan trimestres con variables reportadas, lo que resulta en una muestra de 14.850 empresas. En la construcción de este *dataframe* se utiliza la función *rolling* para obtener las cifras TTM¹⁶ del beneficio neto ('niclean_ltm') y del EBITDA ('ebitda_ltm'), sumando la variable cuatro periodos hacia atrás.

Se genera entonces un segundo *dataset*¹⁷ con la información de mercado, es decir, los precios diarios a cierre de mercado de las empresas. Mientras que la frecuencia de la información es mayor (diaria respecto a trimestral), el número de variables es mucho menos, siendo necesario únicamente obtener el código identificador de la compañía, la fecha, el precio diario y el número de acciones ('gvkey', 'datadate', 'prccd', 'cshoc'). También aquí se obtenía la información en *chunks* para leer el archivo de forma manejable, pero al haberse

¹⁵ En el Anexo 1 se recoge el código (*Compustat cleaning*) por el que se construye este dataset con las variables sobre las empresas.

¹⁶ En el código se ha llamado LTM (*Last Twelve Months*), que es otra forma de referirse a las cifras TTM.

¹⁷ En el Anexo 2 se recoge el código (*Prices cleaning*) por el que se construye este dataset con las variables sobre precios diarios de las acciones.

reducido el número de columnas (sólo cuatro variables), los *chunks* son de 500.000 filas cada uno. El periodo temporal para el que se obtienen precios diarios es de unos 13 años, desde el 4 de enero de 2010 hasta el 29 de diciembre de 2023.

En un primer momento, se obtiene información de precios sobre 31.655 empresas. Se cruza esta información con la del dataset anterior de 14.850 empresas con datos contables suficientes y ordenados, y la muestra combinada resulta ser de 8.713 compañías. Se procede entonces a eliminar: duplicidades (por el código de empresa o por fecha), datos inconsistentes (diferentes precios o número de acciones en una fecha determinada) y “huecos” de información (necesitan tener datos de precio durante el periodo de estudio). Finalmente, la muestra de empresas de las que tenemos información trimestral sobre las variables contables relevantes e información sobre precios diarios, resulta ser de 5.670 elementos. De estas empresas se calcula la capitalización bursátil diaria multiplicando las variables precio por acción y número de acciones.

3.5. Construcción de los múltiplos

Una vez tenemos los dos *datasets*, el de variables TTM calculados a partir de cifras reportadas por la empresa ('compustat_clean_short.csv'), y de la capitalización bursátil diaria de las acciones ('mktcap_clean.csv'), se procede a fusionarlos con el fin de calcular los múltiplos de valoración, siguiendo la definición de dichos múltiplos de la sección anterior¹⁸.

En primer lugar, se eliminan las columnas con las variables que ya no van a utilizarse. A continuación se ordenan los datos en función de las fechas, como se explica a continuación. Es necesario señalar que a la hora de construir los múltiplos de valoración previamente definidos, mientras que el numerador, al depender del precio de cotización, cambiará diariamente, el denominador (y parte de los datos del numerador) se mantienen constantes durante un trimestre, y sólo cambiarán con un nuevo anuncio de resultados trimestrales por parte de la empresa, que cambiaría las variables TTM. Es importante en este sentido ajustar las fechas con cuidado, diferenciando el periodo al que hace referencia la variable reportada, y la fecha en la que ocurre la comunicación o reporte de la información (*filing date*). Es decir, el beneficio por ejemplo del primer trimestre se cierra a 31 de marzo, pero los estados financieros trimestrales no se registran con la autoridad de mercado y por tanto se anuncian públicamente hasta un tiempo después (supongamos en algún momento a finales del mes de

¹⁸ En el Anexo 3 se recoge el código ('*Multiples*') por el que se fusionan los dos datasets y se calculan los múltiplos de valoración.

abril). Es importante tenerlo en cuenta, ya que los múltiplos de valoración continuarán con las cifras TTM que se habían calculado el trimestre anterior hasta que los resultados sean públicos. En el ejemplo anterior, los múltiplos de valoración continuarían construyéndose con las cifras a cierre de año natural (los cuatro trimestres del año anterior) variando únicamente el precio diario, hasta la fecha a finales de abril en la que el anuncio de los resultados trimestrales produciría un ajuste y se recalcularían las variables TTM con los tres últimos trimestres del año anterior y el primero del presente año. Es decir, aunque las cifras que se añaden están cerradas a 31 de marzo, no afectarían al múltiplo de valoración hasta el *filig date* a finales de abril. A partir de ese momento, las variables contables permanecerían sin cambios hasta el *filig date* de julio en que se anunciarían resultados a 30 de junio. La variable 'rdq' se refiere a la fecha de publicación de resultados y es la que se tiene en cuenta para alinear temporalmente los datos y que sean consistentes.

Tras identificar y corregir datos anómalos que pueden distorsionar el análisis (en este caso múltiplos con valores negativos, que se ajustan a cero), se une el dataset de múltiplos con los datos identificativos de las empresas y sus códigos de industria. El resultado es un fichero CSV con múltiplos diarios para 5.679 empresas con datos identificativos que van a permitir ordenarlas por industria ('multiples_clean.csv')

3.6. Cálculo de las discrepancias

Finalmente, una vez calculados los múltiplos de valoración PER y EV/EBITDA para nuestra muestra de empresas, se van a calcular lo que hemos denominado “discrepancias” en tales múltiplos¹⁹. Una discrepancia tendría lugar cuando una empresa está aparentemente sobrevalorada respecto a sus comparables con uno de los dos múltiplos considerados, e infravalorada respecto al otro.

Para identificar la existencia o no de estas discrepancias, es necesario en primer lugar situar a cada compañía en un grupo que incluya a las empresas que se consideran más comparables²⁰. En este caso se agrupan por industrias según el código SIC de dos dígitos al que ya se ha hecho referencia.

¹⁹ En el Anexo 4 se recoge el código ('Cálculo discrepancias') en el que se utiliza el 'multiples_clean.csv' para calcular las discrepancias en los múltiplos de valoración.

²⁰ Ya se ha hecho mención a que este requisito de “comparabilidad” puede ser el talón de Aquiles de una valoración relativa, puesto que si las empresas no son tan similares (y, como se suele decir en el mercado “*there is no such thing as two similar companies*”), falla la premisa de partida y por tanto el resultado no va a ser fiable.

Tras limpiar aquellas observaciones sin SIC o con múltiplos “infinitos” (al tener un denominador de valor cero), se procede a calcular los cuartiles de los múltiplos por cada industria, para cada día. El cuartil es un concepto estadístico que ordena los valores de un conjunto de datos en cuatro partes iguales. Los cuartiles nos ayudan a entender cómo se distribuyen los datos dentro de un conjunto. Normalmente se identifican tres cuartiles específicos (tres puntos de corte que dividen el conjunto en cuatro partes, “cuartiles”): el primer (Q1), el segundo(Q2), y el tercer cuartil (Q3). El primero (Q1) engloba el 25% de los datos inferiores del conjunto, el segundo (Q2), que coincide con la mediana, divide el conjunto en mitades iguales, y el tercero (Q3) abarca hasta el 75% de los datos, dejando el 25% superior por encima. Esta división en cuartiles se realiza para los dos múltiplos (PER y EV/EBITDA) cada día dentro de cada industria. Así, tendríamos que la empresa X, en una fecha determinada tiene por ejemplo un PER en el cuartil 1 (que agrupa los múltiplos más bajos) y un EV/EBITDA en el cuartil 2 (agrupa los múltiplos que son más altos que los del Q1 pero más bajos que la mediana). A cada una de las 5.679 empresas se le asignó, para cada día, un valor del 1 al 4 en las columnas “EV_EBITDA_Quartile” y “PE_Quartile”.

Finalmente, ya tenemos los datos para poder identificar y medir las discrepancias en los múltiplos, realizando la resta de los cuartiles en los que están los ratios en una fecha determinada. Concretamente, la discrepancia se calcula como el cuartil del EV/EBITDA menos el cuartil del PER, apareciendo el resultado en una nueva columna “Diff”. Si la diferencia es cero, ambos múltiplos están en el mismo cuartil y la valoración sería coherente. Si es positiva, es que el cuartil del EV/EBITDA es superior, es decir, el múltiplo sería alto respecto a sus comparables en términos relativos al del PER, que sería bajo respecto a sus comparables en términos relativos al EV/EBITDA (supongamos por ejemplo que el EV/EBITDA estaría en un Q4 y el PER en un Q2), por lo que la acción parecería estar “barata respecto al PER y cara con el EV/EBITDA”. En cambio, si la diferencia es negativa, la posición de los cuartiles sería la contraria, y la acción parecería estar “cara con el PER y barata con el EV/EBITDA”. Si la diferencia (positiva o negativa) es de valor 1, es que la diferencia es de tan solo un cuartil, mientras que si es de 2 o de 3, es que la discrepancia es más pronunciada y los múltiplos se sitúan a dos o incluso tres cuartiles de distancia uno del otro.

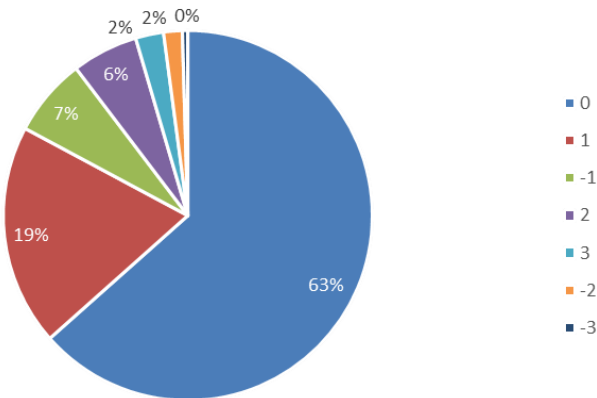
CAPÍTULO IV: RESULTADOS E INTERPRETACIÓN

4.1. Descripción de resultados

A continuación se describirán brevemente los resultados. Con el fin de realizar la interpretación de los mismos, el análisis se centrará en las 10 industrias que más empresas tienen dentro de nuestro *dataset* final, buscando sacar conclusiones que estén respaldadas por un número significativo de empresas.

En una primera aproximación, los resultados generales, sin agrupar por industria, nos muestran que en la mayoría de observaciones (63%) no existen discrepancias entre los múltiplos de valoración y ambos sitúan a las empresas en el mismo cuartil, creando una diferencia de cuartiles de 0. Respecto a las observaciones con discrepancias, podemos ver que el 26% de se sitúan a 1 cuartil de diferencia, siendo más frecuente que las discrepancias sean positivas, es decir, que las empresas estén más baratas respecto al múltiplo PER y caras si se considera el EV/EBITDA. Las observaciones a 2 cuartiles de diferencia son bastante menos, siendo también más predominantes aquellas con discrepancias positivas y un múltiplo de valoración más alto de EV/EBITDA. Esta tendencia se mantiene también para los valores con discrepancias a 3 cuartiles de diferencia, que representan una cantidad escasa de las observaciones totales.

Tabla 6: Descripción de las discrepancias



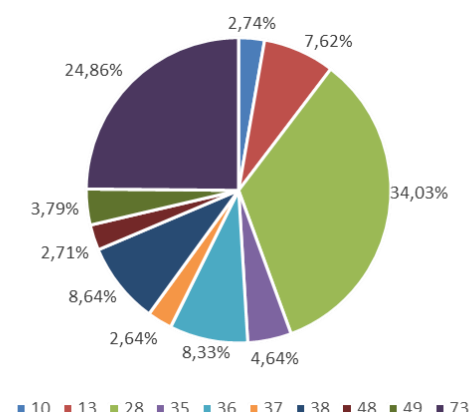
Fuente: elaboración propia

Así pues, aunque ambos múltiplos valoran a las empresas de forma similar respecto a sus comparables (clara mayoría de casos sin discrepancia), es cierto que para un volumen importante de observaciones no se cumple esta regla y debemos intentar darle una explicación aterrizando el análisis, como se ha dicho anteriormente, para casos concretos en una selección de industrias.

Para elegir las 10 industrias, hay que señalar que se han excluido los sectores financieros. Las industrias con código SIC de dos cifras 60 (*Bancos - Cajas de ahorro*) y 67 (*Sociedades de Cartera - Holdings*), pese a tener un número suficiente de empresas y de observaciones para incluirse en el análisis, no serán estudiadas ya que la naturaleza operativa de la industria y la financiera pueden confundirse, resultando en una cuenta de pérdidas y ganancias singular. Por ende, el uso de los múltiplos de valoración en cuestión puede no ser indicativo de una discrepancia, ya que como se ha mencionado anteriormente, si bien el PER es un múltiplo utilizado habitualmente en este tipo de empresas, el EV/EBITDA no se calcula para entidades financieras. La tabla 7 muestra las industrias seleccionadas para el análisis son las siguientes:

Tabla 7: Industrias con mayor número de empresas

SIC	Industria	Nº Empresas	%
10	MINERALES METÁLICOS	81	2,74%
13	PETRÓLEO Y GAS NATURAL	225	7,62%
28	PRODUCTOS QUÍMICOS	1005	34,03%
35	MAQUINARIA	137	4,64%
36	MAQUINARIA ELÉCTRICA Y ELECTRÓNICA	246	8,33%
37	EQUIPOS DE TRANSPORTES	78	2,64%
38	INSTRUMENTOS DE MEDIDA, ANÁLISIS, CONTROL Y OTROS	255	8,64%
48	COMUNICACIONES	80	2,71%
49	SERVICIOS DE AGUA, ELECTRICIDAD, GAS Y SANITARIOS	112	3,79%
73	SERVICIOS COMERCIALES	734	24,86%
Total		2953	100%



Fuente: elaboración propia

Todas estas industrias comprenden un número considerable de empresas, suficiente como para realizar un análisis de las discrepancias en profundidad y tratar de sacar conclusiones. Estas 2.953 empresas de las 10 industrias seleccionadas suponen el 52% del total del *dataset*, por lo que se aprovecha gran parte de la información disponible para el análisis. Las discrepancias de las 10 industrias seleccionadas se recogen en la tabla 8.

Tabla 8: Frecuencia de discrepancias por industria

#	10	13	28	35	36	37	38	48	49	73
0	121579	178915	1442188	127653	276439	92761	350689	57847	76494	674435
1	4981	91299	20317	74808	276439	30014	47912	47574	57012	329065
-1	1037	8365	9763	22115	26967	11548	26549	7460	20372	37850
2	27	36032	180	22115	45730	7613	7725	18185	14196	109433
-2	0	865	21	4588	2622	1838	67	972	5915	4041
3	0	9035	9	9769	1124	1356	2	10455	12163	11571
-3	0	5	3	332	0	0	0	0	1265	0
Total	127624	324516	1472481	261380	629321	145130	432944	142493	187417	1166395

Fuente: elaboración propia

Como se puede observar, en línea con los resultados generales, la mayor parte de las diferencias calculadas en todas las industrias no arrojan discrepancia y la valoración de ambos múltiplos coloca a las empresas en el mismo cuartil. También se observa que el porcentaje de discrepancias con 3 cuartiles de diferencia es escaso respecto al total, y que en todos los casos son más frecuentes las discrepancias positivas (*cheap on PER, expensive on EV/EBITDA*) que las negativas (situación contraria: cara según el PER, barata según el EV/EBITDA).

Entrando a comentar industrias concretas, saltan a la vista algunas en las que la gran mayoría de observaciones no muestran discrepancias: Minería de metales (código SIC 10) y Productos químicos (código SIC 28), aunque es necesario señalar que para esta última industria solo se disponen datos para los 5 primeros años, por lo que resulta más interesante ver la tabla 9 con porcentajes. Por el contrario, otras industrias tienen más de la mitad de los datos con algún tipo de discrepancia: Maquinaria industrial y equipos informáticos (código SIC 35), Maquinaria y componentes eléctricos y electrónicos (SIC 36), Comunicaciones (SIC 48) y Servicios de agua, electricidad, gas y sanitarios (SIC 49). En la siguiente sección se examinará la naturaleza de las industrias para tratar de explicar el comportamiento mencionado atendiendo a sus características de negocio, operativas y financieras.

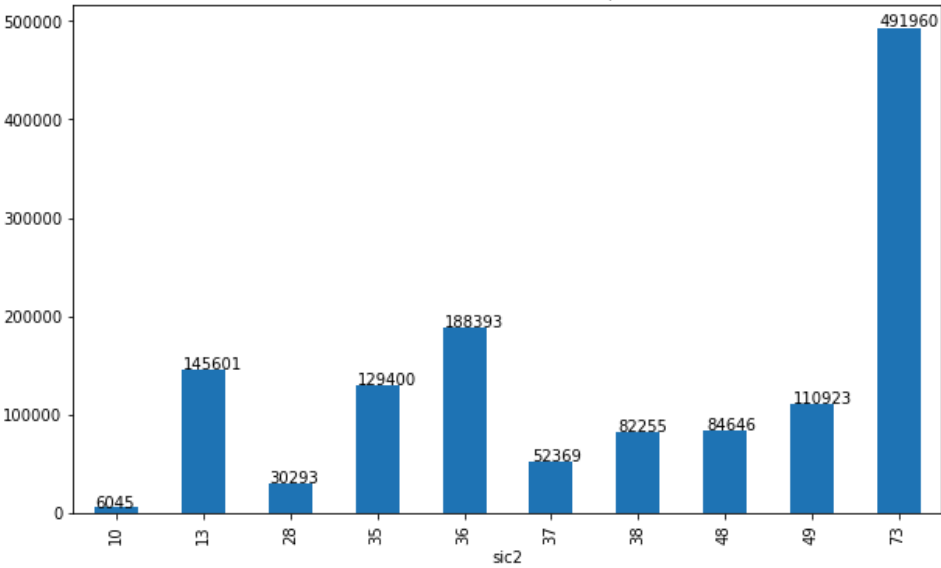
Tabla 9: Porcentaje de discrepancias por industria

%	10	13	28	35	36	37	38	48	49	73
0	95%	55%	98%	49%	44%	64%	81%	41%	41%	58%
1	4%	28%	1%	29%	44%	21%	11%	33%	30%	28%
-1	1%	3%	1%	8%	4%	8%	6%	5%	11%	3%
2	0%	11%	0%	8%	7%	5%	2%	13%	8%	9%
-2	0%	0%	0%	2%	0%	1%	0%	1%	3%	0%
3	0%	3%	0%	4%	0%	1%	0%	7%	6%	1%
-3	0%	0%	0%	0%	0%	0%	0%	0%	1%	0%
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

Fuente: elaboración propia

Si nos fijamos únicamente en los datos que presentan discrepancias para las industrias elegidas (tomando valores de discrepancia de 1, 2, 3, -1, -2, o -3) vemos la siguiente distribución de frecuencia absoluta. Sobresale notablemente la industria de Servicios Comerciales (SIC 73), en parte también debido a que el grupo tiene una gran cantidad de empresas en la muestra.

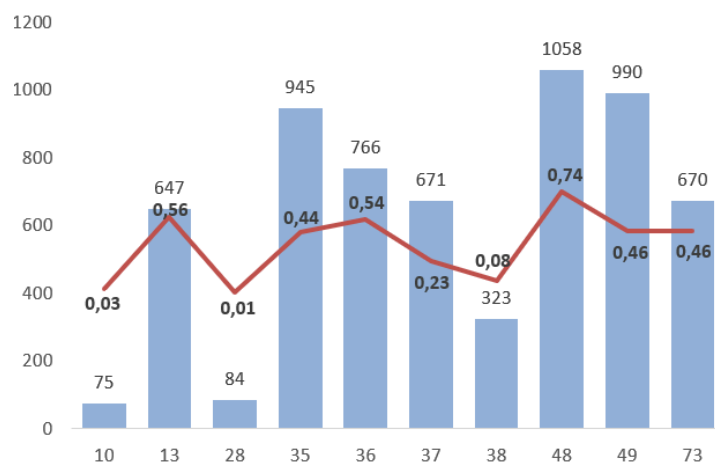
Figura 8: Distribución de discrepancias por industria



Fuente: elaboración propia

Para tratar de valorar el número de discrepancias de cada industria sin que se vea afectado por el número de empresas de cada grupo, calcularemos una media de discrepancia por empresa en cada industria (n° de discrepancias/ n° empresas en cada industria), resultados que se pueden ver en la figura 9. De nuevo, es importante recordar que la industria 28 (Productos Químicos) solo tiene observaciones para 5 años, por lo que el valor medio de 84 discrepancias está calculado proyectando la tendencia disponible para el periodo de años que tienen el resto de industrias.

Figura 9: Discrepancia media por empresa e industria. Índice de discrepancia media



Índice									
10	13	28	35	36	37	38	48	49	73
0,03x	0,56x	0,01x	0,44x	0,54x	0,23x	0,08x	0,74x	0,46x	0,46x

Fuente: elaboración propia

El gráfico muestra también (línea roja) un índice de las discrepancias, que aparece también en la tabla inferior de la gráfica, calculado como una media ponderada de la frecuencia de cada observación, tipo de discrepancia (valores 1,-1, 2,-2, 3 y -3) o no discrepancia (valor 0) por industria. El hecho de que la mayoría de las discrepancias sean positivas le da al índice un valor positivo para todas las industrias, mientras que el gran número de observaciones sin discrepancia hace que el valor sea siempre menor de 1.

Como podemos ver, todas las empresas tienen un índice positivo, lo que nos lleva a la conclusión de que, de forma general, es más frecuente que la empresa aparezca sobrevalorada según el múltiplo EV/EBITDA que según el PER. En base a este índice, se procederá a analizar las industrias concretas para tratar de dar una explicación a esta información.

4.2. Interpretación de resultados. Análisis de las industrias²¹

En esta sección realizamos una reflexión sobre cómo las características de las industrias analizadas pueden haber afectado a los múltiplos de valoración y a las discrepancias. Se presentan en primer lugar consideraciones generales comunes a todos los sectores, para después describir las características básicas de cada uno intentando relacionarlas con las discrepancias.

Entre los ajustes que reconcilian el EV/EBITDA con el PER, tal y como señalan el ya mencionado trabajo de Jullens (2012), podemos destacar: (1) las dotaciones a las amortizaciones (tanto de activos tangibles como intangibles), que están relacionadas con las necesidades de inversión o *capex*, que Jullens (2012) denomina *depreciation adjustment factor* y *amortisation adjustment factor* (D&A) y (2) el nivel de apalancamiento financiero o endeudamiento (*gearing factor*).

Todas las industrias seleccionadas tienen en común que son industrias intensivas en capital (intensidad media-alta), lo que conlleva un alto valor en los activos, inversiones importantes en *capex* y dotaciones elevadas en D&A, aunque estas dependerán en ocasiones de la vida media de dichos activos y de las políticas de amortización. Estas industrias, además, tienen niveles medio-alto de apalancamiento financiero, lo que es habitual en industrias intensivas en capital y con muchos activos fijos. Este mayor nivel de deuda lleva aparejado unos gastos financieros que disminuirán el beneficio neto (a pesar del ahorro fiscal o *tax shield*), aunque también hay que recordar que para la empresa es más barato financiarse con deuda que con *equity*.

Teniendo esto en cuenta, parece normal que las empresas de estas industrias presenten en media valores mayores en el EV/EBITDA que en el PER (discrepancias positivas):

Empresas intensivas en capital, con *capex* alto, y elevados niveles de deuda, presentarán un EV (numerador) elevado en relación al EBITDA (denominador), del que excluyen el D&A y los intereses financieros. Al reconciliar el múltiplo añadiendo las amortizaciones, los intereses y los impuestos para obtener el beneficio neto (denominador) y ajustando la deuda neta y otros factores para dejar el *Equity* en el numerador, puede surgir la discrepancia,

²¹ La breve descripción de la actividad en cada una de las industrias se ha extraído de la Occupational Safety and Health Administration (US Department of Labor) en <https://www.osha.gov/data/sic-manual>

especialmente si el nivel de deuda es alto pero el coste de la misma es bajo, lo que en empresas industriales es habitual.

A continuación vamos a describir la actividad de cada uno de estos grupos industriales. Además de las características ya comentadas, y que se aplican a todas las industrias analizadas, se señalará algún aspecto específico que también pueda influir en la discrepancia:

1. SIC 10 – Minería de metales:

El grupo principal SIC 10 incluye establecimientos dedicados principalmente a la minería, desarrollo de minas o exploración de minerales metálicos. Estas minas son valoradas principalmente por los metales que contienen, los cuales se extraen para ser utilizados como tales o como componentes de aleaciones, productos químicos, pigmentos etc. Se incluyen todas las operaciones realizadas en molinos que trituran, muelen, lavan, secan, sintetizan o calcinan minerales metálicos, o realizan operaciones de separación por gravedad o flotación. Se excluyen fundiciones y refinerías.

Las empresas de este sector muestran discrepancias positivas, esto es múltiplos EV/EBITDA mayores que múltiplos PER, pero con un índice de 0,03, uno de los menores de las industrias analizadas.

Se trata de un sector cíclico, con alta volatilidad y oscilaciones del beneficio neto en función de los precios de las materias primas y otros insumos y *commodities*, lo que provoca que los inversores en acciones exijan unos retornos sobre su inversión más altos y se generen múltiplos PER menores.

2. SIC 13 – Extracción de petróleo y gas natural:

El grupo principal SIC 13 incluye establecimientos dedicados principalmente a la exploración, perforación, operación y mantenimiento de pozos de petróleo y gas, la operación de plantas de gasolina natural y de ciclo, y la gasificación, licuefacción y pirólisis del carbón en el sitio de la mina. Se incluyen actividades básicas como la ruptura de emulsiones y la desarenación del petróleo crudo en la preparación de petróleo y gas que habitualmente se realiza en el sitio del campo, pero no el transporte por tubería de petróleo, gasolina y otros productos derivados.

Las empresas de este sector muestran discrepancias positivas, esto es múltiplos EV/EBITDA mayores que múltiplos PER, pero en este caso con un índice de 0,56, uno de los mayores de las industrias analizadas.

Además de que su intensidad de capital es una de las más altas del conjunto de sectores elegido, se trata de un sector con una volatilidad relativamente alta, que depende del precio del petróleo y del gas natural, y de crecimiento limitado, lo que permite a los accionistas exigir retornos sobre el *equity* (RoE) más altos y genera PER más bajos.

3. SIC 28 – *Productos químicos:*

El grupo principal SIC 28 incluye establecimientos que producen productos químicos básicos o que fabrican productos mediante procesos predominantemente químicos. En general, estas empresas fabrican: (1) productos químicos básicos, como ácidos, álcalis, sales y productos químicos orgánicos; (2) productos químicos para ser utilizados en fabricación adicional, como fibras sintéticas, materiales plásticos, colores secos y pigmentos; y (3) productos químicos terminados para el consumo final, como medicamentos, cosméticos y jabones; o para ser utilizados como materiales o suministros en otras industrias, como pinturas, fertilizantes y explosivos.

Las empresas de este sector muestran discrepancias positivas, esto es múltiplos EV/EBITDA mayores que múltiplos PER, pero con un índice de sólo 0,01, el menor de todas las industrias analizadas.

Se trata además de un sector con fuertes barreras de entrada, algo menos apalancado, con importantes inversiones en I+D, cíclico (con importantes diferencias según el subsector del que se trate), con alta volatilidad y limitado crecimiento, lo que al igual que en los sectores anteriores permite al inversor exigir un alto RoE y produce PER más bajos.

4. SIC 35 – *Maquinaria industrial y equipos informáticos:*

El grupo principal SIC 35 incluye empresas dedicadas a la fabricación de maquinaria y equipo industrial y comercial, y computadoras. Se incluye la fabricación de motores y turbinas; maquinaria para granjas y jardines; maquinaria para construcción, minería y campos petrolíferos; ascensores y equipos de transporte; grúas, monorraíles y camiones y tractores industriales; maquinaria para trabajar metales; maquinaria para industrias específicas; maquinaria industrial general; equipos informáticos y periféricos, y otra maquinaria de

oficina; maquinaria de refrigeración y para la industria de servicios. Se incluyen herramientas de mano motorizadas y máquinas impulsadas por motores integrados o desmontables exceptuando los electrodomésticos.

Las empresas de este sector muestran discrepancias positivas, esto es múltiplos EV/EBITDA mayores que múltiplos PER, y de cierta relevancia, ya que el índice es de 0.44.

Esta categoría empresarial agrupa empresas muy diversas, que varían enormemente en función de las necesidades y tendencias del sector y de los clientes para los que trabajen, si bien en general su beneficio neto presenta un potencial de crecimiento bajo.

5. SIC 36 – Maquinaria y componentes eléctricos y electrónicos:

El grupo principal SIC 36 incluye compañías dedicadas a la fabricación de maquinaria, aparatos y suministros para la generación, almacenamiento, transmisión, transformación y utilización de energía eléctrica. Se incluye la fabricación de equipos de distribución de electricidad; aparatos industriales eléctricos; electrodomésticos; equipos de iluminación y cableado eléctrico; equipos receptores de radio y televisión; equipos de comunicación; componentes y accesorios electrónicos; y otros equipos y suministros eléctricos.

Las empresas de este sector muestran discrepancias positivas, esto es múltiplos EV/EBITDA mayores que múltiplos PER, con un índice de 0,54, uno de los mayores de la muestra de industrias analizadas.

Estas empresas son intensivas en capital de diferente naturaleza y requieren un alto nivel de inversiones (*capex*) tanto en equipos y activos fijos como en I+D o en otros activos intangibles (dado el alto nivel de tecnología necesario), aunque hay cierta variabilidad del beneficio en función de cada subsector y cada tipología de clientes.

6. SIC 37 – Equipos de transportes:

El grupo principal SIC 37 incluye establecimientos dedicados a la fabricación de equipos para el transporte de pasajeros y carga por tierra, aire y agua. Se incluyen vehículos de motor, aeronaves, misiles guiados y vehículos espaciales, barcos, botes, equipo ferroviario y equipo de transporte diverso, como motocicletas, bicicletas y motos de nieve.

Las empresas de este sector muestran discrepancias positivas, esto es múltiplos EV/EBITDA mayores que múltiplos PE, con un índice de 0.23.

Estas empresas, dedicadas a la provisión de equipos de transporte, por carretera o vías férreas entre otros, son también intensivas en capital de diferente naturaleza, y requieren un alto nivel tanto de inversiones (*capex*) como de inventario. Además, en este sector suele existir un elevado nivel de exigencia regulatoria y unas altas barreras de entrada que, si bien en general ofrecen un crecimiento limitado, proporcionan visibilidad futura a los inversores.

7. SIC 38 – Instrumentos de medida, análisis y control:

El grupo principal SIC 38 incluye empresas dedicadas a la fabricación de instrumentos profesionales y científicos para medir, probar, analizar y controlar, y sus sensores y accesorios asociados; instrumentos ópticos y lentes; instrumentos de topografía y dibujo; equipos hidrológicos, hidrográficos, meteorológicos y geofísicos; sistemas y equipos de búsqueda, detección, navegación y guía; instrumentos, equipos y suministros quirúrgicos, médicos y dentales; artículos oftálmicos; equipos y suministros fotográficos y relojes y cronómetros.

Las empresas de este sector muestran discrepancias positivas, esto es múltiplos EV/EBITDA mayores que múltiplos PER, con uno de los índices más bajos de nuestra muestra, de 0,08.

Estas empresas se dedican a la provisión de diferentes instrumentos y equipos pero de menor cuantía financiera y con un grado de intensidad de capital menor que otras industrias de la muestra. También suelen presentar niveles menores de apalancamiento financiero, lo que implica menores gastos financieros, por lo que se generan EV/EBITDA más próximos al PER.

8. SIC 48 – Comunicaciones:

El grupo principal SIC 48 incluye compañías que proporcionan servicios de comunicaciones punto a punto, ya sean de forma auditiva o visual, incluyendo la radiodifusión y televisión. También se incluyen servicios de buscapersonas y bípens, así como aquellos que se dedican al arrendamiento de líneas telefónicas u otros métodos de transmisión telefónica, como líneas de fibra óptica y facilidades de microondas o satélites, y a la reventa del uso de dichos métodos a otros.

Las empresas de este sector muestran discrepancias positivas, esto es múltiplos EV/EBITDA claramente mayores que múltiplos PER, con el mayor índice de todas las industrias analizadas: 0,74.

Son empresas muy intensivas en capital de diferente naturaleza, tanto redes, equipos e inversiones físicas como activos inmateriales (*software*, derechos y patentes). Se trata además de un sector con relativa visibilidad futura, alta regulación y crecimiento incierto, lo que en general genera múltiplos PER bajos en relación al EV.

9. SIC 49 – Servicios de agua, electricidad y sanitarios:

El grupo principal SIC 49 incluye establecimientos dedicados a la generación, transmisión y/o distribución de electricidad, gas o vapor. Estas empresas pueden ser combinaciones de cualquiera de los tres servicios mencionados anteriormente e incluyen servicios, como transporte, comunicaciones y refrigeración. También comprenden sistemas de agua y riego, y sistemas sanitarios dedicados a la recolección y eliminación de basura, aguas residuales y otros desechos mediante la destrucción o procesamiento de materiales.

Las empresas de este sector muestran discrepancias positivas, esto es múltiplos EV/EBITDA mayores que múltiplos PE, con un índice de 0,46, relativamente alto.

Estas empresas son también muy intensivas en capital y necesitan acometer un alto nivel de inversiones en sus redes y equipos de agua o electricidad, moviéndose además en un entorno muy regulado, con alta visibilidad sobre la evolución del negocio, y un crecimiento relativamente bajo, lo que puede generar múltiplos PER relativamente menores.

10. SIC 73 – Servicios comerciales:

El grupo principal SIC 73 incluye empresas dedicadas principalmente a la prestación de servicios no clasificados en otra parte, y establecimientos comerciales que funcionan en base a un contrato o tarifa como publicidad, informes de crédito, cobro de deudas, servicios de correo, reproducción, taquigrafía, agencias de noticias, programación de computadoras, fotocopiado, duplicado, procesamiento de datos, servicios a edificios y servicios de suministro de personal.

Las empresas de este sector también muestran discrepancias positivas, esto es múltiplos EV/EBITDA mayores que múltiplos PE, con un índice relativamente alto, de 0,46.

Esta industria agrupa a empresas de muy diversa índole, y aunque posiblemente su intensidad de capital en activos físicos sea menor, en términos relativos requieren y contabilizan niveles más altos de activos inmateriales (patentes, derechos y otros intangibles). El nivel de apalancamiento financiero también será variable, al ser un sector heterogéneo dependiente de la actividad de cada tipología de cliente, lo que también permite a los diferentes inversores y proveedores de capital exigir un alto nivel de RoE en función de las dinámicas de la industria en cuestión y unos múltiplos PER relativamente más bajos.

CAPÍTULO V: CONCLUSIONES

Los múltiplos de valoración como el ratio PER o el EV/EBITDA son muy utilizados en el día a día de los mercados por distintos profesionales financieros. Estos ratios indican que el precio de una acción parece correcto o es “caro” o “barato” en relación a acciones de empresas similares, lo que puede llevar a decisiones de inversión o ajuste en las carteras de los gestores si creen que se debe a un *mispricing* temporal. Es cierto que empresas líderes o con claras ventajas competitivas pueden consistentemente presentar múltiplos superiores a los de sus comparables sin que ello implique un *mispricing*, sino que merecen un precio (y un múltiplo) superior. Lo mismo puede ocurrir en sentido inverso, cuando una empresa de peor calidad cotiza de forma continuada por debajo de su *benchmark* de comparables. Es por tanto imprescindible conocer bien los fundamentales y la estrategia de la empresa para poder interpretar de forma correcta lo que nos dicen los múltiplos de valoración relativa. Esta interpretación es aún más difícil cuando dichos ratios presentan una contradicción: según un múltiplo el precio de la acción estaría por encima de sus comparables, pero según otro estaría por debajo, por lo que la acción parece estar “*cheap on P/E but expensive on EV/EBITDA*” o a la inversa. Esta contradicción es lo que hemos llamado “discrepancia” y es el objeto de análisis de este TFG. Más concretamente, el TFG se había planteado los siguientes objetivos:

1. Utilizar la base de datos compustat para construir los múltiplos de valoración PER e EV/EBITDA y calcular las posibles discrepancias.

Para la elaboración de este trabajo se ha utilizado la base de datos compustat, de la que se extrajo información sobre variables contables fundamentales y sobre precios diarios de 5.679 empresas durante el periodo 2010-2023. Utilizando lenguaje python se generaron dos datasets de información (uno contable, otro con precios de mercado) que posteriormente se fusionaron para calcular los múltiplos diarios de EV/EBITDA y PER, utilizando información trimestral para construir variables TTM (*Trailing Twelve Months*). Para ello fue necesario diferenciar la fecha de resultados (primer trimestre por ejemplo) de la fecha de comunicación o reporte de dichos resultados, que es algo posterior pero es cuando se conoce la información y se ajustan los múltiplos. Tras agrupar las empresas por industria según la clasificación del código SIC de dos dígitos, se ordenaron estos múltiplos en cuartiles, y se calcularon las discrepancias: diferencias en cuartiles entre la posición del múltiplo EV/EBITDA y la del PER (cuartil donde está el EV/EBITDA - cuartil donde está el PER).

2. Describir las discrepancias encontradas

En un 63% de las observaciones no existen discrepancias entre los múltiplos de valoración situándose ambos en el mismo cuartil (diferencia de cuartiles 0). Un 26% de las discrepancias suponen una diferencia de sólo un cuartil, Respecto a las observaciones con discrepancias, podemos ver que el 26% de se sitúan a 1 cuartil de diferencia, un 8% con 2 cuartiles de diferencia y tan solo un 3% presentan discrepancias con diferencia de 3 cuartiles. En todos los casos es mucho más frecuente que las discrepancias sean positivas, es decir, que las empresas estén baratas respecto al múltiplo PER y caras respecto al EV/EBITDA.

Centrándonos en las 10 industrias con mayor número de observaciones, se observa que en algunas de ellas más de la mitad de las observaciones presentan algún tipo de discrepancia: Maquinaria industrial y equipos informáticos (código SIC 35), Maquinaria y componentes eléctricos y electrónicos (SIC 36), Comunicaciones (SIC 48) y Servicios de agua, electricidad, gas y sanitarios (SIC 49), mientras que en otras, como Minería de metales (código SIC 10) y Productos Químicos (código SIC 28), la gran mayoría de las observaciones no arrojan discrepancias.

Se elaboró un índice de discrepancias calculando la media ponderada de la frecuencia de discrepancias por el tipo de discrepancia (valores 1,-1, 2,-2, 3 y -3) o no discrepancia (valor 0) por industria. Este índice tiene un valor positivo en todos los sectores, lo que era de esperar al ser la mayoría de las discrepancias positivas (EV/EBITDA en un cuartil superior al PER). El gran número de observaciones sin discrepancia hace que el valor sea siempre menor de 1.

3. Reflexionar sobre la mayor o menor frecuencia de discrepancias por industria, tratando de ligar la presencia de las mismas a las características del negocio

Todas las industrias analizadas son intensivas en capital, tienen activos y *enterprise values* de valor alto en los activos, inversiones importantes en *capex* y dotaciones elevadas en D&A. Además, tienen niveles altos de apalancamiento financiero, lo que suele ser común industrias con muchos activos fijos. Este mayor nivel de deuda produce unos gastos financieros que disminuirán el beneficio neto (a pesar del ahorro fiscal ligado a tales intereses). El coste de la financiación vía deuda es más barato que financiarse con *equity*.

Siendo este el caso, pare lógico que las empresas de estas industrias presenten en media valores mayores en el EV/EBITDA que en el PER (discrepancias positivas), ya que empresas intensivas en capital, con *capex* alto, y elevados niveles de deuda, presentarán un EV elevado en relación al EBITDA, que no incluye D&A ni intereses financieros. Al reconciliar el

múltiplo añadiendo las amortizaciones, los intereses y los impuestos para obtener el beneficio neto y ajustando la deuda neta y otros factores para obtener el *equity* en el numerador, puede surgir la discrepancia, especialmente si el nivel de deuda es alto pero el coste de la misma es bajo, o si el valor de los activos es alto pero el D&A es más bajo debido a la vida de los activos o a políticas contables. Estos factores de ajuste y otros se recogen en el trabajo de Jullens (2012)

Para finalizar el TFG, se presenta una línea de investigación para el futuro, que profundizaría en el análisis realizado en este TFG. Sería interesante utilizar las variables de compustat de las empresas de la muestra para calcular numéricamente cada uno de los factores de ajuste de Jullens (2012), con el fin de reconciliar el múltiplo EV/EBITDA con el PER, y poder identificar de forma específica para cada caso qué factor (o factores de ajuste) explican la discrepancia.

Declaración de Uso de Herramientas de Inteligencia Artificial Generativa en Trabajos Fin de Grado

ADVERTENCIA:

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

Por la presente, yo, Carlos de Abajo, estudiante de E-3 Analytics de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado "Análisis de la discrepancia entre los múltiplos de valoración de empresas PER y EV/EBITDA: un enfoque por industria", declaro que he utilizado la herramienta de Inteligencia Artificial Generativa ChatGPT u otras similares de IAG de código sólo en el contexto de las actividades descritas a continuación:

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

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

Fecha: 20/06/2024

Firma: _____

A handwritten signature in black ink, consisting of several loops and a long horizontal stroke at the end, written over a horizontal line.

CAPÍTULO VI: BIBLIOGRAFÍA

Álvarez-Sala Díaz-Caneja, S. (2022): Discrepancias entre EV/EBITDA y PE en la valoración relativa de empresas. *Trabajo Fin de Grado*, Facultad de Ciencias Económicas y Empresariales, Universidad Pontificia Comillas Madrid.

Anderson, K. & Brooks, C. (2006a): Decomposing the price-earnings ratio, *Journal of Asset Management*, vol. 6, pp. 456-469.

Anderson, K. & Brooks, C. (2006b): The long-term price-earnings ratio, *Journal of Business Finance & Accounting*, vol. 33, no. 7-8, pp. 1063-1086.

Ayash, B., Bednarek, Z. & Bordeman, A. (2021): Do Compustat Financial Statement Data Actually Articulate?, *Available at SSRN 3951510*, .

Ball, R., & Brown, P. (2013): An empirical evaluation of accounting income numbers. In *Financial Accounting and Equity Markets*, pp. 27-46. Routledge.

Basu, S. (1977): Investment performance of common stocks in relation to their price-earnings ratios: A test of the efficient market hypothesis , *The journal of Finance*, vol. 32, no. 3, pp. 663-682.

Block, S. (2010): Methods of valuation: Myths vs. reality, *The Journal of Investing*, vol. 19, no. 4, pp. 7-14.

Boettke, P. J. (2010): What happened to "efficient markets"?. *The Independent Review*, 14(3), pp. 363-375.

Bouwens, J., De Kok, T. & Verriest, A. (2019): The prevalence and validity of EBITDA as a performance measure , *Comptabilité-Contrôle-Audit*, vol. 25, no. 1, pp. 55-105.

Chan, R.W. & Lui, B.C. (2011): EV/EBIT ratio: The best of both worlds , *Better Investing*, vol. 60, pp. 27-28.

Chan, R.W. & Lui, B.C. (2010): Gaining insight with the EV/EBITDA ratio , *Better Investing*, vol. 60, no. 3, pp. 27-28.

Damodaran, A. (2012): *Investment valuation: Tools and techniques for determining the value of any asset*, Ed. John Wiley & Sons.

De Bondt, W. F., & Thaler, R. (1985): Does the stock market overreact?. *The Journal of finance*, 40(3), pp. 793-805.

Deng, M., Easton, P.D. & Yeo, J. (2012): Another look at equity and enterprise valuation based on multiples , *Available at SSRN 1462794*, .

Fabozzi, F.J., Focardi, S.M. & Jonas, C. (2017): *Equity Valuation: Science, Art, or Craft?* CFA Institute Research Foundation.

Fama, E. F. (1965): The behavior of stock-market prices. *The journal of Business*, 38(1), pp. 34-105.

Fama, E. F. (1970): Efficient capital markets. *Journal of finance*, 25(2), pp. 383-417.

Foster, E.M. (1970): Price-Earnings Ratio and corporate growth , *Financial Analysts Journal*, vol. 26, no. 1, pp. 96-99.

Fama, E. F., Fisher, L., Jensen, M. C., & Roll, R. (1969): The adjustment of stock prices to new information. *International economic review*, 10(1), pp. 1-21.

Graham, B. & Dodd, D. (1934): *Security Analysis*, Ed. Whittlesey House.

Gompers, P., Kaplan, S.N. & Mukharlyamov, V. (2015): No title , *What do private equity firms say they do?*, .

Goodman, D.A. & Peavy III, J.W. (1983): Industry relative price-earnings ratios as indicators of investment returns , *Financial Analysts Journal*, vol. 39, no. 4, pp. 60-66.

Gordon, M.J. (1959): Dividends, earnings, and stock prices , *The review of economics and statistics*, , pp. 99-105.

Grossman, S. J., & Stiglitz, J. E. (1976): Information and competitive price systems. *The American economic review*, pp. 246-253.

Holland, D.A. (2018): An improved method for valuing mature companies and estimating terminal value , *Journal of Applied Corporate Finance*, vol. 30, no. 1, pp. 70-77.

Jegadeesh, N., & Titman, S. (1993): Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, 48(1), pp. 65-91.

Jensen, M. C. (1978): Some anomalous evidence regarding market efficiency. *Journal of financial economics*, 6(2/3), pp. 95-101.

Jitmaneroj, B. (2017): The impact of dividend policy on price-earnings ratio: The role of conditional and nonlinear relationship , *Review of Accounting and Finance*, vol. 16, no. 1, pp. 125-140.

Jullens, D. (2012). Expensive on P/E; cheap on EV/EBITDA: European Distillers. UBS Investment Research. The footnotes Compendium, pp. 60-65.

Koller, T., Goedhart, M. & Wessels, D. (2015): *Valuation. Measuring and Managing the Value of Companies*, Ed. John Wiley and sons (6th edition).

Lifland, S. (2011): The impact of working capital efficiencies on the enterprise value option: Empirical analysis from the energy sector , *Advances in business research*, vol. 2, no. 1, pp. 57-70.

Liu, J., Nissim, D. & Thomas, J. (2007): Is cash flow king in valuations? , *Financial Analysts Journal*, vol. 63, no. 2, pp. 56-68.

Liu, J., Nissim, D. & Thomas, J. (2002): Equity valuation using multiples , *Journal of Accounting Research*, vol. 40, no. 1, pp. 135-172.

Lo, A. W., & MacKinlay, A. C. (1988): Stock market prices do not follow random walks: Evidence from a simple specification test. *The review of financial studies*, 1(1), pp. 41-66.

Loughran, T. & Wellman, J.W. (2011): New evidence on the relation between the enterprise multiple and average stock returns , *Journal of Financial and Quantitative Analysis*, vol. 46, no. 6, pp. 1629-1650.

Mauboussin, M.J. (2018): What does an EV/EBITDA multiple mean , *Bluemont Investment Research*. Obtenido de:

file:///C:/Users/Mio/Downloads/What-Does-an-EV-EBITDA-Multiple-Mean%20(3).pdf

Miller, M.H. & Modigliani, F. (1961): Dividend policy, growth, and the valuation of shares , *the Journal of Business*, vol. 34, no. 4, pp. 411-433.

- Molodovsky, N. (1995): A theory of price–earnings ratios , *Financial Analysts Journal*, vol. 51, no. 1, pp. 29-43.
- Nicholson, S.F. (1960): Price-earnings ratios , *Financial Analysts Journal*, vol. 16, no. 4, pp. 43-45.
- Nissim, D. (2019): Terminal value , *Columbia Business School Research Paper* , no. 18-12.
- Nissim, D. & Penman, S.H. (2001): Ratio analysis and equity valuation: From research to practice , *Review of Accounting Studies*, vol. 6, pp. 109-154.
- Nyborg, K.G. & Mukhlynina, L. (2016): Survey of Valuation Professionals: Valuation Techniques in Practice , *Swiss Finance Institute* , .
- Ou, J.A. & Penman, S.H. (1989): Accounting measurement, price-earnings ratio, and the information content of security prices , *Journal of accounting research*, vol. 27, pp. 111-144.
- Payne, B., Wong, R. & Tyler, M. (2018): On The Use Of Enterprise Value Multiples As Indicators Of Intrinsic Value In Emerging Markets , *Southwestern Economic Review*, vol. 45, pp. 57-71.
- Penman, S.H. (1996): The articulation of price-earnings ratios and market-to-book ratios and the evaluation of growth , *Journal of accounting research*, vol. 34, no. 2, pp. 235-259.
- Pinto, J. E., Robinson, T. R., & Stowe, J. D. (2019). Equity valuation: A survey of professional practice. *Review of financial economics*, 37(2), 219-233.
- Platt, H., Demirkan, S. & Platt, M. (2010): Free cash flow, enterprise value, and investor caution , *The Journal of Private Equity* , pp. 42-50.
- Rozenbaum, O. (2019): EBITDA and managers' investment and leverage choices, *Contemporary Accounting Research*, vol. 36, no. 1, pp. 513-546.
- Stattman, D. (1980): Book values and stock returns , *The Chicago MBA: A journal of selected papers*, vol. 4, no. 1, pp. 25-45.
- Thorndike Jr, W. N. (2012): *The outsiders: eight unconventional CEOs and their radically rational blueprint for success*. Ed. Harvard Business Press.

Tseng, K.C. (1988): Low price, price-earnings ratio, market value, and abnormal stock returns, *Financial Review*, vol. 23, no. 3, pp. 333-343.

Urbano García, D. (2023): Discrepancias entre EV/EBITDA y PE en la valoración relativa de empresas: un enfoque descriptivo. *Trabajo Fin de Grado*, Facultad de Ciencias Económicas y Empresariales, Universidad Pontificia Comillas Madrid.

Vidal-Garcia, R. & Ribal, J. (2019): Terminal value in SMEs: Testing the multiple EV/EBITDA approach , *Journal of Business Valuation and Economic Loss Analysis*, vol. 14, no. 1, pp. 20180012.

Williams, J.B. (1938): *The Theory of Investment Value*, Ed. Harvard University Press.

Zheng, S.X. & Stangeland, D.A. (2007): IPO underpricing, firm quality, and analyst forecasts, *Financial Management*, vol. 36, no. 2, pp. 1-20.

CAPÍTULO VII: ANEXOS

7.1 Código *Compustat Cleaning*

```
import pandas as pd
##### CORE
# Define el nombre del archivo CSV original y el archivo de salida
#datos_i = 'C:\\Users\\
#datos_f = 'C:\\Users\\
myfolder = "C:\\Users\\Windows\\Downloads\\"
datos_i = myfolder +
"compustat_quarterly_1962-01-2024-02\\compustat_quarterly_1962-01-2024-02.csv"
datos_f = myfolder + "compustat_quarterly_preprocessed.csv"
datos_clean_f = myfolder + "compustat_clean_full.csv"
datos_short_f = myfolder + "compustat_clean_short.csv"
precios_i = myfolder + "compustat_stockprices_daily_2020-2024\\yb8n0w5pwimqua9l.csv"
precios_f = myfolder + "precios.csv"
co_list = myfolder + "company_list.csv"
# Define las columnas que quieres conservar
columnas_seleccionadas = ['gvkey', 'datadate', 'fyearq', 'fqtr', 'cusip', 'conm', 'cheq', 'cshoq', 'dlcq',
                          'dlttq', 'epsfxq', 'ibcomq', 'mibq', 'niq', 'oibdpq', 'pstkq', 'xidoq', 'gind', 'gsubind',
                          'naics', 'rdq', 'sic']
# Define el formato de las columnas
formato_columnas = {'gvkey': 'str', 'datadate': 'str', 'fyearq': 'int', 'fqtr': 'str', 'cusip': 'str', 'conm': 'str',
                    'cheq': 'float', 'cshoq': 'float', 'dlcq': 'float', 'dlttq': 'float', 'epsfxq': 'float', 'ibcomq': 'float',
                    'mibq': 'float', 'niq': 'float', 'oibdpq': 'float', 'pstkq': 'float', 'xidoq': 'float', 'gind': 'str',
                    'gsubind': 'str', 'naics': 'str', 'rdq': 'str', 'sic': 'str'}
# Crea un DataFrame vacío para almacenar los datos filtrados
df_filtrado = pd.DataFrame()
df_filtrado2 = pd.DataFrame()
# Define el tamaño de cada chunk (número de filas por chunk)
chunk = 100000
%%time
##### CORE
```

```

# Crea un dataframe vacío para almacenar los cálculos intermedios
df_mid = pd.DataFrame()
# Procesa cada chunk
for mychunk in pd.read_csv(datos_i, usecols=columnas_seleccionadas, dtype=formato_columnas,
chunks=chunk):
    df_mid.drop(df_mid.index, inplace=True)
    df_mid = mychunk
    # Debt in current liabilities (dlcq) + Long-term debt total (dlttq) +
    # Preferred stock at carrying value (pstkq) + Minority interest Balance Sheet (mibq) -
    # Cash and short term investments (ceqq) is the EV adjustment
    df_mid['EV_adj'] = df_mid.dlcq.fillna(0) + df_mid.dlttq.fillna(0) + df_mid.pstkq.fillna(0)
    + df_mid.mibq.fillna(0) - df_mid.cheq.fillna(0)
    # Concatena el chunk actual al DataFrame general
    df_filtrado = pd.concat([df_filtrado, df_mid], ignore_index=True)
# Guarda el DataFrame final en un nuevo archivo CSV
df_filtrado.to_csv(datos_f, index=False)
print("El archivo ha sido procesado y guardado con éxito.")
print("Archivo con " + str(df_filtrado.shape[0]) + " filas y " + str(df_filtrado.shape[1]) + " columnas")
print("Número de compañías en el archivo: " + str(df_filtrado.gvkey.nunique()))
print("Desde " + str(df_filtrado.datadate.min()) + " hasta " + str(df_filtrado.datadate.max()))
print("Número de observaciones con el ajuste de EV sin definir: " +
str(df_filtrado['EV_adj'].isna().sum()))
df_filtrado.columns
# Compruebo que la creación del dataset se haya realizado correctamente
# archivo_csv= 'C:\\Users\\
df = pd.read_csv(datos_f)
print(df.head())
##### CORE
#Ajusto el formato de las fechas
df_filtrado['datadate'] = pd.to_datetime(df_filtrado['datadate'])
df_filtrado['rdq'] = pd.to_datetime(df_filtrado['rdq'])
df_filtrado.tail()

```

#En contra de lo que parecería razonable, en el dataframe original hay observaciones con los índice duplicados

```
df_filtrado[df_filtrado.duplicated(subset = ['gvkey','datadate'], keep = False)]
```

```
##### CORE
```

#Quitamos estas observaciones con índices duplicados para empresa, año y trimestre

```
df_clean = df_filtrado.reset_index().drop_duplicates(subset = ['gvkey','datadate'], keep = False)
```

#NO Hay observaciones con la fecha vacía

```
df_clean.datadate.isnull().sum()
```

```
##### CORE
```

#Creamos un nuevo multi-índice que incluya para cada empresa TODOS los trimestres posibles,

#No tan solo aquellos para los que hay observaciones como pasaba hasta ahora.

#Para eso vamos a hacer el producto de una lista con todas las empresas y otra con todas las fechas

#de los trimestres.

#Aquí generamos la lista de todas las empresas

```
key_ids = df_clean.gvkey.unique()
```

```
print(len(key_ids))
```

```
print(key_ids)
```

```
##### CORE
```

#Aquí generamos la lista de todos los trimestres entre la primera y última fecha en el fichero

```
q_list = (pd.date_range(pd.to_datetime(df_clean.datadate.min()),
                        pd.to_datetime(df_clean.datadate.max()) + pd.offsets.QuarterBegin(1), freq='Q')
         .strftime('%Y-%m-%d')
         .tolist())
```

```
q_list[0:4]
```

```
%%time
```

```
##### CORE
```

#Creamos un nuevo multi-índice que incluya para cada empresa TODOS los trimestres posibles,

#No tan solo aquellos para los que hay observaciones como pasaba hasta ahora

```
idx = pd.MultiIndex.from_product([key_ids,q_list], names=['gvkey','datadate'])
```

```
idx.size
```

```
idx
```

```
%%time
```

```
##### CORE
```

```

#Indexamos el dataframe de nuevo según empresa y trimestre
df_clean = df_clean.set_index(['gvkey','datadate'])
# y lo reindexamos de manera que ahora cada empresa tiene observaciones (aunque sean vacías)
# para todos los trimestres
df_clean = df_clean.reindex(idx)
df_clean = df_clean.drop(columns = ['index'])
df_clean.tail()
%%time

##### CORE

#Calculo las cifras LTM. Primero agrupo por compañía y luego hago el cálculo utilizando rolling
cuatro periodos atrás

#con un mínimo de cuatro periodos. Como antes me he asegurado de que todas las compañías tienen
filas para todos los periodos

#no tengo que preocuparme por saltos en los datos
df_clean["niclean_ltm"]=df_clean.groupby('gvkey', group_keys=False).ibcomq.apply(
    lambda x: x.rolling(4, min_periods = 4).sum())
df_clean["ebitda_ltm"]=df_clean.groupby('gvkey', group_keys=False).oibdpq.apply(
    lambda x: x.rolling(4, min_periods = 4).sum())
df_clean.iloc[54:64,:]
df_clean.tail()
%%time

##### CORE

# Exporto el dataframe expandido, que comprende un panel balanceado a la fuerza con muchas
observaciones NaN
df_clean.reset_index().to_csv(datos_clean_f, index=False)

```

Proceso de eliminación de casos

No basta con eliminar las observaciones para las que faltan datos.

Hay que mantener sólo aquellas empresas que no tengan "huecos" en sus series temporales (es decir, que no tengan, por ejemplo, datos para el trimestre 1, 2 y 4, pero no el 3, o para un año sí y otro no).

Si no hacemos esto, cuando luego juntemos nuestros números con los precios y extendamos los números contables para pasarlos de trimestrales a diarios, los extenderemos más de lo razonable. En el ejemplo de antes, para la primera empresa, en vez de extender los datos contables del segundo trimestre desde mediados de julio hasta octubre, como no habría datos para el tercer trimestre, extenderíamos el número de junio hasta enero del año siguiente (cuando aparecerían los datos del cuarto trimestre). Esto no es razonable, ya que el hecho de que falte un trimestre en medio de la serie no es "normal".

Para hacer esta limpieza, utilizamos un código que no es inmediatamente fácil de seguir. Así que primero lo aplicamos con este dataset de juguete a modo de ejemplo.

```
# Example dataset
df = pd.DataFrame(data = { 'company':[item for item in ['A','B','C','D'] for i in range(3)],
                          'year': [2000,2001,2002,2001,2002,2003,2000,2001,2002,2001,2002,2003],
                          'value1':pd.Series([5, 10,2,15,20,5,6,2,45], index=[1,2,3,4,5,6,7,9,11]),
                          'value2':pd.Series([3, 6,7,8,6,51,6,4,9], index=[0,1,2,3,4,5,7,9,11])},
                  index = [0,1,2,3,4,5,6,7,8,9,10,11])

df

# Calculo una variable que es NaN si alguna de las dos columnas
# que me interesa es NaN
df.loc[(~df['value1'].isnull() & ~df['value2'].isnull()),'present' ] = True
df

# Ahora busco el primer y ultimo año en el que la variable de interés no es NaN
m = df[df['present'].notnull()].groupby('company').agg(mini = ('year','min'),
                                                       maxi=('year','max')).reset_index()

# Fusiono este pequeño dataframe de vuelta con el original
new = df.merge(m)

new

#Mantengo solamente las observaciones entre la primera y la última de
#tal modo que hay potencialmente periodos enteros para cada empresa, pero
#como se puede ver en el ejemplo, es posible que haya "agujeros".
new = new[(new['year']>=new['mini']) & (new['year']<=new['maxi'])]

new

# Agrupo por empresa y me quedo con aquellas que no tengan NINGUN valor NaN en present
new.groupby('company').filter(lambda x: ~x['present'].isnull().any())

##### CORE

# Ahora aplicamos el código al dataframe entero
# Las columnas que me interesa que estén siempre presentes son las dos
# magnitudes LTM, el ajuste de EV y el filling date
df_clean.loc[(~df_clean['niclean_ltm'].isnull() &
              ~df_clean['ebitda_ltm'].isnull() &
              ~df_clean['EV_adj'].isnull() &
```

```

        ~df_clean['rdq'].isnull()
    ),'present' ] = True
print("Number of observations without all relevant data: " +
      str(df_clean.present.isna().sum()) + ", representing " +
      str(df_clean.present.isna().sum() / df_clean.shape[0]) +
      " of all data.")
print("Leaving " + str(df_clean.shape[0] - df_clean.present.isna().sum())
      + " potentially useful observations.")
%%time
##### CORE
# Ahora busco el primer y ultimo año en el que la variable de interés no es NaN
m = df_clean[df_clean['present'].notnull()].reset_index().groupby('gvkey').agg(mini =
('datadate','min'),
                                maxi=('datadate','max')).reset_index()
# Fusiono este pequeño dataframe de vuelta con el original
new = df_clean.reset_index().merge(m)
new.tail()
##### CORE
#Mantengo solamente las observaciones entre la primera y la última de
#tal modo que hay potencialmente periodos enteros para cada empresa, pero
#como se puede ver en el ejemplo, sigue siendo posible que haya "agujeros"
new = new[(new['datadate']>=new['mini']) & (new['datadate']<=new['maxi'])]
print("Number of observations wth potentially continuous time series " + str(new.shape[0]))
print("Out of this, the number of observations without all relevant data is: " +
      str(new.present.isna().sum()) + ", representing " +
      str(new.present.isna().sum() / new.shape[0]) +
      " of all new data.")
print("Número de empresas en este dataset: " + str(new.gvkey.nunique()))
new.tail()
##### CORE
# Ahora corremos el mismo código de arriba para eliminar empresas
# que no tengan un periodo corrido entero
# Agrupo por empresa y me quedo con aquellas que no tengan NINGUN valor NaN en present

```

```

df_short = new.groupby('gvkey').filter(lambda x: ~x['present'].isnull().any())
print("Number of observations dropped " + str(new.shape[0] - df_short.shape[0]))
print("Number of observations with actually continuous time series " + str(df_short.shape[0]))
print("Out of this, the number of observations without all relevant data is: " +
      str(df_short.present.isna().sum()))
print("Número de empresas en este dataset: " + str(df_short.gvkey.nunique()) + ", " +
      str(new.gvkey.nunique() - df_short.gvkey.nunique()) + " less than before.")
df_short.reset_index().columns
##### CORE
df_out = df_short.loc[:,['gvkey','datadate','rdq','conm','gsubind','naics','sic',
'EV_adj','niclean_ltm','ebitda_ltm','present','mini','maxi']]
df_out.shape
companies = pd.DataFrame(df_out.gvkey.unique())
companies.to_csv(co_list, index = False)
df_out = pd.read_csv(datos_short_f, dtype = {'gvkey':'str'})
df_out.dtypes
df_out.head()
%%time
##### CORE
# Exporto el dataframe reducido sin NaN
df_out.reset_index().to_csv(datos_short_f, index=False)

```

7.2 Código *Prices cleaning*

```

import pandas as pd # Very important to have version 2.0 of pandas
import pandas_market_calendars as mcal
# You many need to pip install pandas-market-calendars first
import datetime as dt
import numpy as np
# Define el nombre del archivo CSV original y el archivo de salida
myfolder = "C:\\Users\\Windows\\Downloads\\"
precios_f = myfolder + "precios_all.csv"

```

```

mktcap_f = myfolder + "mktcap_clean.csv"
# Importo la lista de empresas sobre las que hay datos contables y que, por tanto, me interesan
cos = myfolder + "company_list.csv"
co_list = pd.read_csv(cos, dtype = 'str')
# Defino las columnas que quiero conservar
columnas_seleccionadas = ['gvkey', 'datadate', 'precd', 'cshoc'] #'ajexdi'
columnas_tipo = {'gvkey': str, 'datadate': str, 'precd': float, 'cshoc': float}
# Defino los nombres de los archivos
data_folder = myfolder = "C:\\Users\\Windows\\Downloads\\compustat_daily\\"
archivos = ['daily_20-24.csv', 'daily_15-20.csv', 'daily_10-15.csv']
# Visualizo el dataset de precio para ver las columnas que quiero seleccionar
precios_i = data_folder + archivos[0]
chunk= 100
reader = pd.read_csv(precios_i, dtype=str, iterator=True)
df = reader.get_chunk(chunk)
print(df.columns)
%%time
# Crea un DataFrame vacío para almacenar los datos filtrados
df_p = pd.DataFrame()
# Define el tamaño de cada chunk (número de filas por chunk)
chunk = 500000
i = 0
# Para cada archivo
for myfile in archivos:
    # Procesa cada chunk de un archivo
    precios_i = data_folder + myfile
    for mychunk in pd.read_csv(precios_i, usecols=columnas_seleccionadas, dtype= columnas_tipo,
chunks=chunk):
        i = i+1
        # Concatena el chunk actual al DataFrame general
        df_p= pd.concat([df_p, mychunk], ignore_index=True)
        print(i)
print("El archivo ha sido procesado con éxito.")

```



```

print("El dataframe de precios tiene " + str(df_p.shape[0]) + " líneas")

print("Comprende desde " + str(df_p.datadate.min()) + " hasta " + str(df_p.datadate.max()) + " y
representa "

    + str(df_p.gvkey.nunique()) + " empresas, mientras que tenemos datos contables well-behaved tan
solo de " +

    str(len(co_list)))

# Filtramos todas aquellas compañías que no están en la lista que hemos importado con aquellas para
las que

# tenemos series temporales de datos contables consecutivas

a = list(co_list.iloc[:,0])

dfp_filtrado = df_p[df_p['gvkey'].isin(a)]

print("El número de empresas para las que tenemos datos contables well-behaved y precios en el
intervalo estudiado es: "

    + str(dfp_filtrado.gvkey.nunique()))

%%time

# Tenemos muchas observaciones duplicadas

dfp_filtrado.duplicated(keep = False).sum()

%%time

# Nos deshacemos de los duplicados

dfp_clean = dfp_filtrado.drop_duplicates(keep = 'first')

%%time

# Tenemos más problemas. Hay observaciones de la misma empresa y fecha con distinto numero de
acciones

# o precio, probablemente acciones con más de una clase o ADRs.

# Vamos a borrar cualquier empresa que sufra este problema en cualquier momento

bad = dfp_clean[dfp_clean.duplicated(subset = ['gvkey','datadate'], keep = False)]

bad.shape

dfp_clean["bad"] = dfp_clean.duplicated(subset = ['gvkey','datadate'], keep = False)

%%time

dfp_clean = dfp_clean.set_index(['gvkey','datadate']).sort_index()

dfp_clean[dfp_clean.bad].head()

%%time

# Ahora mantenemos sólo los gvkeys de empresas que no tienen ningún bad

dfp_out = dfp_clean.groupby('gvkey').filter(lambda x: ~x["bad"].any())

print("Quedan " + str(dfp_out.reset_index().gvkey.nunique()) + " compañías, de las " +

```

```

    str(dfp_clean.reset_index().gvkey.nunique()) + " que había antes.")
print("La suma de duplicados es ahora: " + str(dfp_out.bad.sum()))
print("El dataframe tiene ahora " + str(dfp_out.shape[0]) + " filas")
#Calculamos el market cap
dfp_out['mktcap'] = dfp_out['cshoc'] * dfp_out['prccd']
# Vamos a contar cuantas observaciones hay por año para cada empresa, ver cuál es el máximo y
deshacernos de
# las que no tengan suficientes.
dfp_out.head()
%%time
#Lo primero es ver cuántas observaciones hay de máximo para cada año
ni = dfp_out.reset_index()
ni['datadate'] = pd.to_datetime(ni['datadate'])
ni = ni.set_index('datadate')
ni = ni.groupby([ni.index.year,'gvkey']).size().reset_index().rename(columns = {0:'data'})
nito = ni.groupby('datadate').agg(size = ('data','max'), company_index = ('data','idxmax'))
nito
# Ahora vamos a asegurarnos que todas las empresas tienen todas las fechas.
# Para eso primero hacemos pivot del dataframe lo que iguala las fechas para todas las empresas
# y llena de NaN los espacios que no estaban definidos y luego colapsaremos este pivoted dataframe
para
# que tenga la estructura original
df2 = dfp_out.reset_index().loc[:,['datadate','gvkey','mktcap']]
df3 = df2.pivot(index = 'datadate', columns = 'gvkey', values = 'mktcap')
df3.head()
df4 = pd.melt(df3.reset_index(), id_vars = ['datadate'])
df4.head()
#Vemos que, efectivamente, este nuevo dataframe tiene más líneas que el anterior,
#al tener ahora todas las fechas para cada año y no sólo aquellas para las que había datos
#para cada compañía
df4.shape
df4['datadate'] = pd.to_datetime(df4['datadate'])
df4 = df4.set_index('datadate')

```

```

%%time
ni = df4.groupby([df4.index.year,'gvkey']).size().reset_index().rename(columns = {0:'data'})
nito = ni.groupby('datadate').agg(max_size = ('data','max'), min_size = ('data','min'))
#Ahora, todas las empresas tienen el mismo número de fechas
nito

# Vamos a poner un mínimo al número de días en que una empresa cotiza en un año para mantenerla
en el dataframe.

# Más adelante rellenaremos los días que falten con observaciones del último precio válido
min_days = 200

# Hay que contar el número de NaN que tiene cada empresa para cada año y en función de eso, y de en
qué año

# empezó y terminó de cotizar, mantenerla en la muestra o no.
#df4.set_index('datadate')
g1 = df4[df4.value.isnull()]
g2 = g1.groupby(['gvkey',g1.index.year]).size().reset_index().rename(columns={0:'missing'})
g2.head()
m = dfp_out.reset_index().groupby('gvkey').agg(mini = ('datadate','min'), maxi = ('datadate','max'))
m
g3 = g2.merge(m.reset_index(),how='left')
g3['minyr'] = pd.DatetimeIndex(g3['mini']).year
g3['maxyr'] = pd.DatetimeIndex(g3['maxi']).year
g3['out'] = (g3['missing'] > min_days) & (g3['datadate']>g3['minyr']) & (g3['datadate']<g3['maxyr'])
g3.head()
%%time

# Ahora mantenemos sólo los gvkeys de empresas que no tienen ningún out. Es decir, que todas sus
años

# entre el primero y el último para los que había observaciones tienen más de un mínimo de precios
diarios
g4 = g3.groupby('gvkey').filter(lambda x: ~x['out'].any())
print("Quedan " + str(g4.reset_index().gvkey.nunique()) + " compañías, de las " +
      str(dfp_out.reset_index().gvkey.nunique()) + " que había antes.")
df5 = df4[df4.gvkey.isin(g4.gvkey)]
df5 = df5.rename(columns={'value':'mktcap'})
df5.head()

```

```

# Need to update pandas to at least 2.0 for this to work

# El forward fill hace que los precios válidos de hoy se extiendan a mañana, si mañana hay NaN.

# Los NaN al principio de la serie antes de que la empresa cotizase se mantienen

# Los NaN del final, después de que la empresa dejase el mercado en principio se reescribirían con el
último precio

# válido, pero para evitar eso "limit_area = inside" hace que sólo se reescriban NaN entre valores
válidos y no, los últimos

# Esto es un ejemplo

d = {'A': [np.NaN,np.NaN,2,3,np.NaN,4,np.NaN], 'B':[np.NaN,1,2,np.NaN,np.NaN,4,np.NaN]}
ex = pd.DataFrame(index=[2010,2011,2012,2013,2014,2015,2016], data= d)

ex

ex2 = ex.reset_index().melt(id_vars=['index'])

ex2

# Aplicamos el limit_area para que se vea su efecto en una compañía

ans = ex2[ex2['variable'] == 'A'].ffill(limit_area = 'inside')

ans

# Ahora hacemos el ejemplo con dos compañías para que se parezca a nuestro dataset real

# Pero, desgraciadamente, esta opción no está disponible si agrupamos el dataframe, como se ve aquí.
Así que necesitamos

# una estrategia un poco menos evidente

ans2 = ex2.groupby('variable').ffill(limit_area = 'inside')

ans3 = pd.concat([ex2['variable'], #para que se mantenga la columna de variable
                 ex2.groupby('variable').transform(lambda x: x.ffill(limit_area = 'inside'),
                 ], axis = 1)

ans3

%%time

# Ahora, para el dataframe real

df_final = pd.concat([df5['gvkey'],
                    df5.groupby('gvkey').transform(lambda x: x.ffill(limit_area = 'inside'),)
                    ], axis = 1)

df_final.head()

%%time

# Guarda el DataFrame final en un nuevo archivo CSV

df_final.to_csv(mktpcap_f)

```

7.3 Código *Multiples*

```
import pandas as pd

##### CORE

# Define el nombre de los archivos CSV procesados y el archivo de salida
# Mi versión de dónde están los datos
myfolder = "C:\\Users\\Carlos\\Documents\\1.NOCHE\\"
mktcap_f = myfolder + "mktcap_clean.csv"
datos_short_f = myfolder + "compustat_clean_short.csv"
multiples = myfolder + "multiples_clean.csv"

# Define el tamaño de cada chunk (número de filas por chunk)
chunk = 100000

%%time

##### CORE

# Lee los datos de market cap
df_p = pd.read_csv(mktcap_f)
print(df_p.head())

##### CORE

# Lee los datos contables
df_i = pd.read_csv(datos_short_f)
print(df_i.head())

# Selecciono las columnas relevantes del dataframe de datos contables
df_c = df_i.loc[:,['gvkey','rdq','EV_adj','niclean_ltm','ebitda_ltm']]

# Creo un dataframe con los datos identificativos
df_n = df_i.loc[:,['gvkey','conm','sic','naics','gsubind']]
df_n = df_n.drop_duplicates(subset = ['gvkey'], keep = 'last')
df_n.shape
df_n.head()

print("Missing values in sic in price dataframe: " + str(df_n.sic.isna().sum()))

# Comprobar que no hay NaN en rdq ni en datadate
print("Missing values in datadate in price dataframe: " + str(df_p.datadate.isna().sum()))
print("Missing values in filling date in accounting dataframe: " + str(df_c.rdq.isna().sum()))
```

```

%%time
# Convertir esas columnas a fecha
df_p['datadate'] = pd.to_datetime(df_p['datadate'])
df_c['rdq'] = pd.to_datetime(df_c['rdq'])
# Elijo los índices
df_c.set_index(['gvkey', 'rdq'], inplace=True)
df_c.index = df_c.index.set_names(['gvkey', 'date'])
df_p.set_index(['gvkey', 'datadate'], inplace=True)
df_p.index = df_p.index.set_names(['gvkey', 'date'])
# Selecciono un buen ejemplo para comprobar el merge.
# Un gvkey que esté en los dos sets, que tenga cuentas para 4Q2009 que se publiquen pronto en enero
del 2010 y pueda ver
# que el merge da precios en enero y nada de contabilidad hasta el día de la publicación.
# Encontramos los gvkeys que estan en los dos datasets y vemos que uno empieza la contabilidad a
finales de 2017
a = set(df_p.reset_index().gvkey.unique())
b = set(df_c.reset_index().gvkey.unique())
c = list(a.intersection(b))
c[0:4]
df_c.loc[(32775,slice(None)),:].head(10)
df_p.loc[(32775,slice('2017-12-01','2017-12-20')),:]
# Así que cogemos el gvkey de arriba en estas fechas como test para la fusión
%%time
# Hago el merge de los datasets
df_merged = df_p.merge(df_c, how='left', left_index=True, right_index=True)
df_merged.to_pickle("C:\\Users\\Carlos\\Documents\\1.NOCHE\\df_merged.pkl")
#Desde aqui puedes leer importando df_merged
print("El dataframe de precios tenía " + str(df_p.shape[0]) + " observaciones y el dataframe fusionado
tiene "
      + str(df_merged.shape[0]))
df_merged.head()
# Chequeo con el ejemplo que he visto arriba
df_merged.loc[(32775,slice('2017-12-01','2017-12-20')),:]
# Y el siguiente reporte

```

```

df_merged.loc[(32775,slice('2018-03-10','2018-03-21')),:]
# Propagar hacia adelante los datos contables
# No nos preocupa propagar el último cuarto de datos demasiado lejos
# Porque primero vamos a eliminar las observaciones sin precio
df_multiples = df_merged[df_merged.mktcap.notnull()]
print("Observaciones del nuevo dataframe: " + str(df_multiples.shape[0]))
print("Hemos eliminado " + str(df_merged.shape[0] - df_multiples.shape[0]) + " observaciones")
print("Quedan " + str(df_multiples['mktcap'].isnull().sum()) + " observaciones con precios NaN")
# Ahora podemos propagar la información contable hacia delante, pero
# SIEMPRE dentro de la misma empresa, es decir, agrupando por gvkey
df_multiples = df_multiples.groupby('gvkey').ffill()
#Hago el calculo del EV y los múltiplos
df_multiples['EV'] = df_multiples['mktcap'] + df_multiples['EV_adj']
df_multiples['PE'] = df_multiples['mktcap'] / df_multiples['niclean_ltm']
df_multiples['EV_EBITDA'] = df_multiples['EV'] / df_multiples['ebitda_ltm']
##### ESTO ES DISCUTIBLE
# Algo hay que hacer con los múltiplos negativos, se pueden convertir en NaN y luego desacerarnos de
la empresa entera
# o solo de la observación, o se pueden considerar 0.
# Esto último es lo que voy a hacer
print("Había " + str(sum(df_multiples['PE']<0)) + " ratios P/E (un " + str(sum(df_multiples['PE']<0) /
df_multiples.shape[0] *100) +
"% del total) y " + str(sum(df_multiples['EV_EBITDA']<0)) + " ratios EV/EBITDA negativos (un
" +
str(sum(df_multiples['EV_EBITDA']<0) / df_multiples.shape[0] *100) + "% del total)")
#Corregimos esos casos
df_multiples.loc[df_multiples['PE']<0,'PE'] = 0
df_multiples.loc[df_multiples['EV_EBITDA']<0,'EV_EBITDA'] = 0
##### HABRÍA QUE TRATAR LOS MÚLTIPLOS DEMASIADO ALTOS
# Fusiono el dataframe con los datos de las empresas para poder agruparlas por industria
print("El dataset de múltiplos tenía " + str(df_multiples.shape[0]) +" observaciones antes de la fusión")
df_n = df_n.set_index('gvkey')
df_multiples = df_multiples.reset_index(level = 1).merge(df_n, how = 'left', left_index = True,
right_index = True)

```

```

print("y ahora tiene " + str(df_multiples.shape[0]))
df_multiples = df_multiples.reset_index().set_index(['gvkey','date'])
df_multiples.tail()
# Borro las observaciones para las que, habiendo precio, no hay datos contables
print("Había " + str(df_multiples['EV_adj'].isnull().sum()) + " observaciones con precio, pero sin datos
contables")
df_multiples = df_multiples[df_multiples['EV_adj'].notnull()]
print("Ahora hay " + str(df_multiples.shape[0]) + " observaciones")
%%time
# Exporto el dataframe
output_file_path = 'C:\\Users\\Carlos\\Documents\\1.NOCHE\\df_multiples.csv'
df_multiples.to_csv(output_file_path, index=True)

```

7.4 Código Cálculo discrepancias

```

import pandas as pd
myfolder = "C:\\Users\\Carlos\\Documents\\1.NOCHE\\"
multiples = myfolder + "df_multiples.csv"
df = pd.read_csv(multiples)
print(df.head())
df.head()
%%time
# Convertir esas columnas a fecha
df['date'] = pd.to_datetime(df['date'])
print(df.columns)
# Defino la función para calcular cuartiles
def calculate_quartile(group, column):
    group[column + '_Quartile'] = pd.qcut(group[column], 4, labels=False, duplicates='drop') + 1
    return group
import numpy as np
df['sic'].dtype
df['sic2']=np.floor(df['sic']/100).astype(int)
print(df.head())

```



```

#Compruebo si hay algun sic na
print("Había " + str(df['sic'].isnull().sum()) + " observaciones sin sic")
inf_values_EV_EBITDA = np.isinf(df['EV_EBITDA']).sum()
print("Había " + str(inf_values_EV_EBITDA) + " observaciones con EV_EBITDA infinito")
df = df[~np.isinf(df['EV_EBITDA'])]
print("Ahora hay " + str(df.shape[0]) + " observaciones")
infinite_values_PE = np.isinf(df['PE']).sum()
print("Había " + str(infinite_values_PE) + " observaciones con PE infinito")
df = df[~np.isinf(df['PE'])]
print("Ahora hay " + str(df.shape[0]) + " observaciones")

%%time
df = df.groupby(['sic2','date']).apply(lambda x: calculate_quartile(x, 'EV_EBITDA'))
df = df.groupby(['sic2','date']).apply(lambda x: calculate_quartile(x, 'PE'))
df.head()
df['Diff'] = df['EV_EBITDA_Quartile'] - df['PE_Quartile']
print(df.head())
df[df['Diff']!= 0].shape[0]
df[df['Diff']== 0].shape[0]
df[df['Diff']> 0].shape[0]
df[df['Diff']< 0].shape[0]
df.to_pickle("C:\\Users\\Carlos\\Documents\\1.NOCHE\\df_discrepancias.pkl")

# Descripcion datos
#Ver para un sic las diffs
condicion1=df['sic2']==60
count1=df.loc[condicion1,'Diff'].value_counts()
print(count1)
condicionA = df['Diff'] != 0
condicionB = df['sic2'] == 81
df_filtradoA = df.loc[condicionA & condicionB]
frecuencias = df_filtradoA['Diff'].value_counts()
print(frecuencias)
suma_frecuencias = frecuencias.sum()
print("La suma de las frecuencias es:", suma_frecuencias)

```

```

condicionAA = df['Diff'] != 0
# Filtrar el DataFrame según la condición general
df_filtradoAA = df.loc[condicionAA]
# Agrupar por 'sic2' y contar las frecuencias de 'Diff' en el DataFrame filtrado
frecuencias_por_sic2 = df_filtradoAA.groupby('sic2')['Diff'].value_counts()
# Mostrar las frecuencias por grupo de 'sic2'
print("Frecuencias de 'Diff' por cada valor de 'sic2' con la condición Diff != 0:")
print(frecuencias_por_sic2)
# Sumar las frecuencias por cada grupo de 'sic2'
suma_frecuencias_por_sic2 = frecuencias_por_sic2.groupby(level=0).sum()
# Mostrar la suma de las frecuencias por cada valor de 'sic2'
print("La suma de las frecuencias por cada valor de 'sic2' es:")
print(suma_frecuencias_por_sic2)
import matplotlib.pyplot as plt
plt.figure(figsize=(10, 6)) # Ajustar el tamaño de la figura
suma_frecuencias_por_sic2.plot(kind='bar')
plt.xlabel('sic2')
plt.ylabel('Suma de Frecuencias de Diff')
plt.title('Suma de Frecuencias de Diff por sic2')
plt.show()

```