



Universidad  
Francisco de Vitoria  
**UFV** Madrid

---

# Trabajo de Fin de Grado

## Variables Macroeconómicas y su Influencia en el Tejido Empresarial Español: Un Estudio de Constituciones y Disoluciones

### **Trabajo Fin de Grado**

Grado en Análisis de Negocios – Business Analytics  
Mayo de 2024

Autor:

**Ignacio López de Carrizosa Grosso**

Tutor:

**Prof. Dra. Ana Lazcano de Rojas**

**Facultad de Facultad de Derecho, Empresa y Gobierno**

Universidad Francisco de Vitoria



*"No todo lo que se cuenta cuenta, y no todo lo que cuenta se cuenta."*

- Albert Einstein

## *Resumen*

Este Trabajo de Fin de Grado explora cómo distintas variables macroeconómicas influyen el sistema empresarial español, con un enfoque particular en las constituciones y disoluciones empresariales. La investigación se concentra en la identificación de factores clave que impactan la estabilidad y el desarrollo de las empresas en España, utilizando modelos de regresión para analizar y prever tendencias dentro del mercado empresarial. Además, se busca informar y democratizar el acceso a los hallazgos del estudio, contribuyendo al debate público y a la toma de decisiones estratégicas. Este enfoque se alinea con la promoción de un entorno empresarial resiliente, con miras a fortalecer la economía española en un contexto global cambiante.

## *Palabras Clave*

Regresión lineal múltiple y simple, variables macroeconómicas, disoluciones y constituciones empresariales, ecosistema empresarial español, modelo.

## *Abstract*

This Final Degree Project explores how different macroeconomic variables influence the Spanish business system, with a particular focus on business incorporations and dissolutions. The research concentrates on identifying key factors that impact the stability and development of firms in Spain, using regression models to analyse and forecast trends within the business market. In addition, it seeks to inform and democratise access to the study's findings, contributing to public debate and strategic decision-making. This approach is aligned with the promotion of a resilient business environment, with a view to strengthening the Spanish economy in a changing global context.

## *Keywords*

Multiple and simple linear regression, macroeconomic variables, business dissolutions and incorporations, Spanish business ecosystem, model.

## RESUMEN AMPLIADO

Este Trabajo de Fin de Grado (TFG) comenzó con un interés en investigar cómo la pandemia del COVID-19 afectó al sistema empresarial español. Sin embargo, el enfoque evolucionó hacia un análisis más amplio y diversificado de cómo ciertas variables macroeconómicas influyen en el sistema empresarial español.

La primera parte del TFG, titulada "Ingeniería del Dato", se centró en el meticuloso proceso de Extracción, Transformación y Carga (ETL) de datos relevantes para el estudio. Este proceso fue fundamental para asegurar la calidad y utilidad de los datos para el análisis subsiguiente, permitiendo una comprensión profunda y precisa del impacto de la COVID-19 en el tejido empresarial español, con datos obtenidos de fuentes fiables como el Instituto Nacional de Estadística (INE) y epdata.

En la transición hacia la segunda parte del trabajo, titulada "Análisis del Dato", el TFG tomó un giro hacia un enfoque más generalizado, analizando la variabilidad del sistema empresarial español bajo la influencia de variables macroeconómicas. Este cambio refleja una adaptación y evolución del objetivo inicial, abarcando un espectro más amplio de análisis que no solo considera el impacto directo de la pandemia sino también cómo el entorno económico y las políticas gubernamentales modelan el panorama empresarial español.

Este cambio de enfoque permitió una exploración detallada de las dinámicas empresariales, revelando complejas interacciones entre la creación de empresas, el capital desembolsado y variables clave como el PIB y el IPC. A través de técnicas de análisis de datos avanzadas, se identificaron patrones y correlaciones significativas, proporcionando insights valiosos sobre la estabilidad y adaptabilidad del tejido empresarial en España.

El resumen narrativo integrado abarca las observaciones y hallazgos de ambos enfoques, destacando la importancia de considerar un espectro amplio de factores cuando se examina el impacto en el sistema empresarial. La combinación de un enfoque inicial centrado en la pandemia, seguido de un análisis más amplio de variables macroeconómicas, ofrece una comprensión más completa y matizada de las dinámicas empresariales en España.

El trabajo final culmina con recomendaciones basadas en el análisis realizado. Este TFG no solo contribuye al cuerpo académico sobre las disoluciones empresariales en España, sino que también proporciona herramientas analíticas que pueden ser aplicadas en futuras investigaciones y prácticas empresariales.

## Índice de Contenidos

---

<b>INTRODUCCIÓN .....</b>	<b>12</b>
<i>Descripción del Problema .....</i>	<i>12</i>
<i>Motivación del Trabajo.....</i>	<i>12</i>
<i>Organización de la Memoria.....</i>	<i>12</i>
<i>Estado del Arte.....</i>	<i>13</i>
<b>OBJETIVOS Y EVOLUCIÓN .....</b>	<b>18</b>
<i>Evolución .....</i>	<i>18</i>
<i>Objetivos Planteados.....</i>	<i>20</i>
<b>PRIMERA PARTE: INGENIERÍA DEL DATO .....</b>	<b>21</b>
<b>1. Introducción a la ingeniería del dato .....</b>	<b>21</b>
<b>2. Origen de los Datos.....</b>	<b>21</b>
<b>3. Descripción de los Datos .....</b>	<b>21</b>
3.1. <i>Número total de observaciones y variables en los datasets (limpias).....</i>	<i>21</i>
3.2. <i>Formato de los datos en bruto y cualquier transformación aplicada. ....</i>	<i>22</i>
3.4. <i>Periodicidad y temporalidad de los datos. ....</i>	<i>25</i>
<b>4. Justificación de la Elección de Variables .....</b>	<b>26</b>
4.1. <i>Explicación de la selección de variables.....</i>	<i>26</i>
4.2. <i>Relación con los objetivos del TFG. ....</i>	<i>26</i>
<b>5. Análisis Exploratorio de Datos .....</b>	<b>27</b>
5.1. <i>Descripción de medidas de tendencia central y dispersión para las variables cuantitativas. ....</i>	<i>27</i>
5.1.2 <i>Medidas de tendencia central .....</i>	<i>27</i>
5.1.3 <i>Medidas de dispersión.....</i>	<i>27</i>
<b>6. Gráficos Descriptivos .....</b>	<b>29</b>
6.1. <i>Gráficos sobre variables y relaciones.....</i>	<i>29</i>

6.2. Patrones, tendencias o correlaciones observadas en los gráficos. ....	30
<b>SEGUNDA PARTE: ANÁLISIS DEL DATO .....</b>	<b>32</b>
<b>8. Introducción al análisis del dato .....</b>	<b>32</b>
<b>9. Selección y Preparación de Datos para el Análisis .....</b>	<b>32</b>
<b>10. Modelos Analíticos: Desarrollo y Aplicación (Regresiones Lineales) .....</b>	<b>34</b>
10.1. Modelo Analítico Supervisado (Regresiones Lineales Simples) .....	34
Explicación del Modelo: .....	34
Proceso de Selección de Variables: .....	36
Desarrollo de modelos: .....	36
Evaluación del Modelo con Medidas de Error/Precisión Específicas: .....	37
10.2. Modelo Analítico Supervisado (Regresiones Lineales Múltiples) .....	39
Explicación del Modelo: .....	39
10.2.1 PRIMERA REGRESIÓN LINEAL MÚLTIPLE (IPC y PIB) .....	40
Proceso de Selección de Variables .....	40
Desarrollo del Modelo .....	40
Evaluación del Modelo con Medidas de Error/Precisión Específicas: .....	41
10.2.2 SEGUNDA REGRESIÓN LINEAL MÚLTIPLE (Variables Macroeconómicas) ...	41
Proceso de Selección de Variables .....	41
Desarrollo del Modelo .....	42
Reconstrucción del modelo en base a la multicolinealidad .....	42
Evaluación del Modelo con Medidas de Error/Precisión Específicas .....	44
10.2.2 TERCERA REGRESIÓN LINEAL MÚLTIPLE (Variables más amplias) .....	44
Proceso de Selección de Variables .....	44
Desarrollo del Modelo .....	45
Evaluación del Modelo con Medidas de Error/Precisión Específicas .....	46
<b>11. Medidas de Adecuación de los Modelos .....</b>	<b>47</b>
11.1. Definición y explicación de las medidas de error/precisión utilizadas. ....	47

11.2. Comparación de los resultados obtenidos en los modelos. ....	49
11.2.1 Medidas de la regresión.....	49
11.2.2 Medidas de Error .....	51
11.3 Aplicación de Pruebas Estadísticas para la Comparación de Modelos.....	52
11.3.1 Prueba de Wilcoxon .....	52
11.3.2 Prueba de Friedman.....	54
<b>12. Visualización de Datos y Resultados de Modelos .....</b>	<b>55</b>
12.1. Gráficos de Dispersión .....	55
12.2. Gráficos de Residuos .....	57
<b>13. Explicación de Resultados.....</b>	<b>58</b>
13.1. Explicación comprensiva de los resultados de los modelos.....	58
13.2. Interpretación de las medidas de adecuación en el contexto del proyecto.....	60
<b>14. Conclusiones .....</b>	<b>62</b>
<b>15. Resumen Narrativo.....</b>	<b>63</b>
<b>TERCERA PARTE: ANÁLISIS DE NEGOCIO .....</b>	<b>64</b>
<b>16. Introducción al Análisis de Negocio .....</b>	<b>64</b>
<b>17. Conclusiones .....</b>	<b>64</b>
17.1 Evaluación y Respuestas los Objetivos Planteados .....	64
<b>18. Opiniones Profesionales .....</b>	<b>66</b>
18.1 Resumen de la Respuestas Obtenidas .....	67
18.3 Insights obtenidos de las preguntas .....	68
<b>19. Recomendaciones y Propuestas .....</b>	<b>69</b>
19.1 Recomendaciones .....	69
19.1.1 Empresarios, Ejecutivos y Emprendedores.....	69
19.1.2 Formuladores de Políticas .....	70
19.2 Propuestas .....	71
19.2.1 Tipo de Propuesta .....	71



19.2.2 Problema y Solución de la Propuesta .....	72
19.2.3 Explicación en Profundidad.....	73
<b>20. Resumen Narrativo.....</b>	<b>74</b>
<b>21. Bibliografía .....</b>	<b>76</b>
<b>ANEXOS .....</b>	<b>78</b>
<b>Anexo 1: Tablas .....</b>	<b>78</b>
<b>Anexo 2: Ilustraciones .....</b>	<b>82</b>
<b>Anexo 3: Entrevistas Completas .....</b>	<b>85</b>
<i>Información de las Entrevistas.....</i>	<i>85</i>
<i>Entrevista Alberto de Santos .....</i>	<i>86</i>
<i>Entrevista José María Martín-Mateos .....</i>	<i>88</i>

## Índice de Figuras

---

<b>Ilustración 1.</b> Captura de pantalla de los diagramas de dispersión correspondientes a las evoluciones de las sociedades Constituidas y disueltas en el sector de la hostelería desde el año 2012 al 2022. ....	29
<b>Ilustración 2.</b> Captura de pantalla del diagrama de cajas correspondiente a la distribución del valor indicador “ICEA” según la comunidad autónoma.....	30
<b>Ilustración 3.</b> Captura de pantalla del histograma correspondiente a la evolución del número de sociedades disueltas voluntariamente en 2021 en todas las comunidades autónomas españolas. ....	30
<b>Ilustración 4.</b> Resultados del test de Wilcoxon para las disoluciones empresariales.....	53
<b>Ilustración 5.</b> Resultados del test de Wilcoxon para las constituciones empresariales.....	53
<b>Ilustración 6.</b> Resultados del test de Friedman para disoluciones empresariales .....	54
<b>Ilustración 7.</b> Resultados del test de Friedman para constituciones empresariales .....	54
<b>Ilustración 8.</b> Análisis de importancia de las variables mediante Random Forest, indicando el peso de cada variable en la predicción de constituciones empresariales.....	56
<b>Ilustración 9.</b> Captura de pantalla del diagrama de dispersión correspondiente a la evolución de sociedades de forma jurídica “S.L.” constituidas en Madrid desde el año 2000 al 2022.....	82
<b>Ilustración 10.</b> Gráfico de la matriz de dispersión para variables del modelo 3.....	82
<b>Ilustración 11.</b> Gráfico de dispersión de variables específicas, en el contexto de constituciones empresariales. ....	83
<b>Ilustración 12.</b> Gráfico de residuos correspondiente al modelo1.....	83
<b>Ilustración 13.</b> Gráfico de residuos correspondiente al modelo2.....	84
<b>Ilustración 14.</b> Gráfico de residuos correspondiente al modelo3.....	84

## Índice de Tablas

<b>Tabla 1.</b> Variables de LIMP.BBDD 1 clasificadas por tipo de variable e información sobre ellas.....	22
<b>Tabla 2.</b> Variables de LIMP.BBDD 2 clasificadas por tipo de variable e información sobre ellas.....	22
<b>Tabla 3.</b> Medidas de tendencia central de la primera base de datos.....	27
<b>Tabla 4.</b> Medidas de dispersión de la primera base de datos. ....	28
<b>Tabla 5.</b> Tabla comparativa de resultados para los modelos de disoluciones.....	49
<b>Tabla 6.</b> Tabla comparativa de resultados para los modelos de constituciones.....	50
<b>Tabla 7.</b> Tabla comparativa de medidas de error para los modelos de disoluciones.....	51
<b>Tabla 8.</b> Tabla comparativa de medidas de error para los modelos de constituciones.....	52
<b>Tabla 9.</b> Variables de LIMP.BBDD 3 clasificadas por tipo de variable e información sobre ellas.....	78
<b>Tabla 10.</b> Variables de LIMP.BBDD 5 clasificadas por tipo de variable e información sobre ellas.....	78
<b>Tabla 11.</b> Variables de LIMP.BBDD 6 y 7 clasificadas por tipo de variable e información sobre ellas .....	79
<b>Tabla 12.</b> Variables de LIMP.BBDD 8 clasificadas por tipo de variable e información sobre ellas.....	79
<b>Tabla 13.</b> Medidas de Normalidad .....	79
<b>Tabla 14.</b> Medidas de Normalidad .....	80
<b>Tabla 15.</b> Medidas de tendencia central de la segunda base de datos. ....	80
<b>Tabla 16.</b> Medidas de tendencia central de la tercera base de datos.....	80
<b>Tabla 17.</b> Medidas de tendencia central de la cuarta base de datos. ....	80
<b>Tabla 18.</b> Medidas de tendencia central de la quinta base de datos. ....	80
<b>Tabla 19.</b> Medidas de tendencia central de la sexta y séptima base de datos.....	80
<b>Tabla 20.</b> Medidas de dispersión de la segunda base de datos.....	81
<b>Tabla 21.</b> Medidas de dispersión de la tercera base de datos. ....	81
<b>Tabla 22.</b> Medidas de dispersión de la cuarta base de datos. ....	81
<b>Tabla 23.</b> Medidas de dispersión de la quinta base de datos.....	81
<b>Tabla 24.</b> Medidas de dispersión de la sexta base de datos.....	81

# INTRODUCCIÓN

## Descripción del Problema

Este Trabajo de Fin de Grado se centra en explorar el comportamiento del sistema empresarial español, con un interés particular en las constituciones y disoluciones empresariales. El problema de investigación radica en entender cómo diversas variables macroeconómicas pueden influir en estas dinámicas empresariales, lo cual es crucial para identificar patrones y factores de riesgo o éxito en el contexto empresarial actual. El estudio busca investigar el impacto que estos factores pueden tener en la estabilidad y el desarrollo de las empresas, proporcionando una base para futuras políticas y estrategias que mejoren la salud empresarial en España.

## Motivación del Trabajo

Este Trabajo de Fin de Grado se inició para comprender los efectos de la pandemia de COVID-19 en el tejido empresarial español. Originalmente centrado en el impacto directo de la pandemia, el estudio se amplió para abordar las dinámicas subyacentes de las constituciones y disoluciones empresariales en España debido a la limitación de los datos específicos de la pandemia.

El proyecto evolucionó hacia el análisis de cómo diversas variables macroeconómicas, como el PIB y el ICEA, influyen en el panorama empresarial. Este enfoque más amplio permite una exploración más rica y diversificada de los factores que afectan la estabilidad y sostenibilidad empresarial, tanto en condiciones normales como en crisis.

La expansión del estudio refleja un compromiso con una comprensión más integral de las condiciones empresariales, abordando la influencia de factores económicos y gubernamentales en las dinámicas empresariales. Los hallazgos pretenden enriquecer la literatura académica y ofrecer orientación para políticas y estrategias que mejoren la resiliencia y adaptabilidad del sector empresarial español.

## Organización de la Memoria

La memoria de este Trabajo de Fin de Grado se estructura en tres partes esenciales, dedicadas a distintas fases del análisis del sistema empresarial español en relación con variables macroeconómicas y su impacto en las constituciones y disoluciones empresariales. En la primera sección se describe el proceso de Extracción, Transformación y Carga (ETL) de los datos, fundamental para asegurar la calidad y relevancia de la

información utilizada en estudios subsiguientes. Se abordan las metodologías de selección de variables y el almacenamiento seguro de datos, elementos que constituyen la base para análisis estadísticos avanzados.

La segunda parte del documento se centra en la implementación de técnicas de análisis para evaluar el efecto de las variables macroeconómicas sobre la dinámica empresarial. Se detalla la elección de modelos analíticos y se discute su eficacia a través de diversas métricas estadísticas, subrayando la importancia de las visualizaciones para facilitar la comprensión de las complejas interacciones que afectan al sector.

La sección final amalgama los insights y datos recabados para ofrecer recomendaciones estratégicas y proporcionar perspectivas prácticas. Este capítulo sintetiza análisis técnico y consideraciones de negocio, reflejando cómo los resultados pueden influir en decisiones políticas y empresariales. Se incluye también la interacción con expertos del sector para enriquecer y validar el análisis, culminando con propuestas concretas destinadas a promover la estabilidad y el crecimiento del ecosistema empresarial.

Cada sección contribuye a una comprensión integral del tema, asegurando que la memoria cumple con los estándares académicos mientras aborda aplicaciones prácticas relevantes para el contexto empresarial contemporáneo.

## **Estado del Arte**

La evolución del tejido empresarial en España constituye un fascinante campo de estudio, especialmente al considerar cómo ha respondido y se ha adaptado a una serie de contextos críticos a lo largo de su historia. Este análisis no solo revela la resiliencia y la capacidad de adaptación de las empresas españolas, sino que también proporciona una ventana a la comprensión de la dinámica económica y social del país en diferentes épocas. A lo largo de los siglos, desde la industrialización en el siglo XIX hasta los desafíos contemporáneos como la crisis financiera global y la pandemia de COVID-19, el tejido empresarial español ha experimentado transformaciones significativas que reflejan tanto las influencias internas como las externas.

En el siglo XIX, España se embarcó en un camino hacia la industrialización, marcando un cambio decisivo en su estructura económica y empresarial caracterizado por la adopción de innovaciones tecnológicas y cambios socioeconómicos y políticos profundos. Este proceso no fue homogéneo ni lineal, presentando variaciones regionales y sectoriales notables.

Uno de los estudios clave en este ámbito es el de Nadal (1975), quien analiza cómo la industrialización en España se vio influenciada por factores como la disponibilidad de

recursos naturales, la infraestructura de transporte y las políticas gubernamentales. A diferencia de otros países europeos, la industrialización en España comenzó más tarde y avanzó a un ritmo más lento. Esto se debió en parte a la estructura agraria del país, la cual retrasó el desarrollo de un sector industrial robusto.

La revolución industrial en España también estuvo marcada por importantes innovaciones tecnológicas, especialmente en la industria textil y metalúrgica. Tortella (1994) destaca cómo la introducción de nuevas máquinas y procesos de producción transformó estas industrias, aunque la adopción de tecnología fue desigual en diferentes regiones del país. Por ejemplo, Cataluña se convirtió en un centro importante para la industria textil, mientras que el País Vasco lideró en la producción de hierro y acero. Además, la industrialización estuvo estrechamente ligada a cambios socioeconómicos significativos. Ringrose (1996) señala que el crecimiento de la industria trajo consigo un cambio en la distribución de la población, con un movimiento creciente de personas del campo a las ciudades en busca de trabajo en las fábricas. Esto no solo cambió el paisaje urbano de España, sino que también tuvo un impacto relevante en la estructura social y económica del país.

En el ámbito político, la industrialización coincidió con un periodo de inestabilidad y cambios. Carr (1966) argumenta que las tensiones políticas del siglo XIX, incluyendo las guerras carlistas y los conflictos entre liberales y conservadores, tuvieron un impacto directo en el desarrollo económico del país. Las políticas gubernamentales en áreas como el comercio, la inversión extranjera y la regulación industrial jugaron un papel crucial en la configuración del proceso de industrialización.

En resumen, la industrialización del siglo XIX en España fue un proceso complejo influenciado por una variedad de factores tecnológicos, socioeconómicos y políticos. Aunque su desarrollo fue más lento y menos uniforme en comparación con otros países europeos, sentó las bases para el crecimiento económico y los cambios sociales en los siglos siguientes.

El siglo XX, sin embargo, presentó desafíos únicos, especialmente con la Guerra Civil Española y la posterior dictadura de Franco. Estos eventos no solo alteraron la estructura económica del país, sino que también impusieron políticas y restricciones que tuvieron un impacto duradero en las empresas españolas.

Durante la Guerra Civil, la economía española sufrió severamente debido a la destrucción de infraestructuras y la interrupción de las actividades comerciales y productivas. Preston (1996) describió cómo la guerra desencadenó una crisis económica aguda, caracterizada por la escasez de materias primas y bienes de consumo, así como por

la disrupción de los mercados. Las empresas tuvieron que enfrentarse a enormes desafíos para mantener sus operaciones en un entorno tan inestable.

Tras el conflicto, la dictadura de Franco implementó un régimen autárquico que buscaba la autosuficiencia económica de España. Este modelo, como señaló Viñas (1984), se caracterizó por una fuerte intervención estatal en la economía, incluyendo la nacionalización de industrias clave y la imposición de estrictos controles sobre la producción, el comercio y los precios. Estas políticas limitaron significativamente la capacidad de las empresas para operar en un mercado libre y restringieron su acceso a mercados internacionales.

La autarquía también condujo a la creación de un complejo sistema de licencias y permisos que afectó a todas las áreas de la actividad empresarial. Este sistema burocrático no solo obstaculizó el crecimiento empresarial, sino que también fomentó la corrupción y el favoritismo. La falta de competencia y la protección a ciertas industrias formaron un entorno empresarial ineficiente y poco competitivo. Payne (1993) destaca cómo la industria pesada, como la minería y la siderurgia, recibió un apoyo considerable, mientras que sectores como la agricultura y los servicios fueron relativamente descuidados. Esto llevó a desequilibrios en el desarrollo económico y agravó las disparidades regionales en España.

La dictadura también tuvo un impacto significativo en las relaciones laborales. La supresión de sindicatos y la imposición de una estricta regulación laboral limitaron los derechos de los trabajadores y contribuyeron a un clima de tensión y descontento en el ámbito laboral.

En conclusión, la Guerra Civil y la dictadura de Franco tuvieron un impacto importante y prolongado en el tejido empresarial español de manera que las políticas económicas y las restricciones comerciales impuestas durante este periodo no solo restringieron el crecimiento y la competitividad de las empresas, sino que también moldearon de manera significativa la estructura económica y social de España.

La transición de España a la democracia en la década de 1970 y su posterior integración en la Unión Europea (UE) en 1986 marcaron un período de transformación económica y empresarial significativa. Estos cambios, fundamentales en la historia reciente de España, impulsaron una liberalización económica y una apertura hacia mercados internacionales, creando un impacto profundo en la expansión y diversificación de las empresas españolas.

La democratización trajo consigo una serie de reformas económicas que buscaban modernizar la economía española. Estas reformas incluyeron la desregulación de los

mercados, la privatización de empresas estatales y la liberalización de los flujos de capital. Guillén (1997) argumenta que este nuevo entorno económico fomentó un clima más propicio para el emprendimiento, la innovación y la inversión extranjera, lo que fue crucial para el desarrollo empresarial en España.

La integración en la UE fue un factor clave en este proceso de transformación. Como señala Holman (1996), la adhesión a la Unión no solo significó el acceso a un mercado mucho más amplio, sino también la adopción de políticas y normativas comunes europeas. Esto incentivó a las empresas españolas a adaptarse a estándares más competitivos, mejorando su eficiencia y capacidad para competir en el mercado global. Sectores como el turismo, la construcción, la banca y las telecomunicaciones experimentaron un crecimiento notable, y las empresas españolas comenzaron a consolidarse en mercados internacionales.

A su vez, este período de transformación también presentó desafíos. La rápida integración en la economía europea y global expuso a las empresas españolas a una competencia más intensa, requiriendo ajustes significativos en términos de calidad, innovación y eficiencia operativa. Royo (2000) destaca que la transición económica llevó a cambios en el mercado laboral, exigiendo una fuerza laboral más cualificada y flexible, lo que a su vez planteó desafíos en términos de formación y adaptación de los trabajadores.

Finalizando esta etapa, se puede decir que la transición a la democracia y la integración en la UE marcaron el inicio de una era de modernización y expansión para el tejido empresarial español. Estos cambios no solo transformaron la estructura económica del país, sino que también establecieron las bases para su crecimiento y desarrollo en las décadas siguientes.

La siguiente etapa que destacar es la crisis financiera global de 2008, la cual tuvo un impacto profundo y duradero en la economía española, afectando particularmente a las pequeñas y medianas empresas (PYMES). Este período crítico se caracterizó por una serie de desafíos económicos, incluyendo una mayor vulnerabilidad financiera, disminución de la rentabilidad y un aumento significativo en los niveles de endeudamiento de las empresas.

La crisis, originada por el colapso del mercado inmobiliario y la crisis de liquidez en los mercados financieros internacionales, rápidamente se propagó a la economía real, afectando a numerosos sectores. En España, las PYMES, que constituyen una parte significativa del tejido empresarial, se vieron especialmente afectadas. Estas empresas enfrentaron dificultades para acceder a financiamiento debido a la restricción crediticia impuesta por las entidades bancarias, que se encontraban en una situación de



incertidumbre y riesgo. Según Molina y Mora (2010), la restricción del crédito fue uno de los factores clave que exacerbó la vulnerabilidad de las PYMES durante la crisis.

Además, la crisis financiera llevó a una disminución significativa en la demanda tanto a nivel nacional como internacional, lo que resultó en una reducción de las ventas y la rentabilidad de las empresas. Como señala Saurina (2009), esta situación se vio agravada por la dependencia de muchas PYMES españolas del mercado interno, que experimentó una contracción considerable. La combinación de una menor demanda y dificultades de financiamiento puso a muchas empresas en una posición precaria, llevando a un aumento en las tasas de insolvencia y cierre de negocios. El endeudamiento de las PYMES se convirtió en un problema crítico durante este período. Muchas empresas que habían tomado préstamos durante los años de bonanza económica se encontraron con dificultades para cumplir con sus obligaciones financieras en un contexto de recesión. Según Ruesga y Da Silva (2011), el alto nivel de endeudamiento limitó la capacidad de las empresas para invertir y adaptarse a las nuevas condiciones del mercado, restringiendo su potencial de recuperación y crecimiento.

La crisis financiera de 2008 también tuvo implicaciones a largo plazo en el panorama empresarial español. La necesidad de reestructuración y adaptación llevó a muchas empresas a revisar sus modelos de negocio, buscar nuevos mercados y mejorar su eficiencia operativa. Este proceso de adaptación no fue fácil, pero fue crucial para la supervivencia y el futuro crecimiento de las PYMES en España.

A fin de cuentas, esta crisis representó un período de desafío significativo para las PYMES en España, marcado por una mayor vulnerabilidad financiera, disminución de la rentabilidad y un aumento en los niveles de endeudamiento. Este evento no solo afectó la economía a corto plazo, sino que también tuvo un impacto perdurable en la forma en que las empresas operan y se adaptan a un entorno económico cambiante.

El último acontecimiento importante en el que se han llevado a cabo estudios sobre el tema a tratar habla sobre la pandemia de COVID-19, que comenzó a principios de 2020, y que ha tenido un impacto sin precedentes en el tejido empresarial global, incluyendo a España. Este evento disruptivo ha afectado a empresas de todos los tamaños y sectores, destacando la importancia crítica de la innovación, la adaptabilidad y el apoyo gubernamental para la supervivencia y recuperación empresarial.

La pandemia trajo consigo desafíos inmediatos y severos para las empresas, incluyendo interrupciones en las cadenas de suministro, cierres obligatorios de negocios no esenciales y una caída drástica en la demanda de consumidores y empresas. En España, donde, como se cita anteriormente, el tejido empresarial está compuesto en gran medida por

PYMES, el impacto fue particularmente agudo. Según un estudio de González et al. (2020), muchas de estas empresas enfrentaron dificultades financieras significativas debido a la disminución de ingresos y la incertidumbre del mercado.

En respuesta a estos desafíos, la innovación y la adaptabilidad se convirtieron en factores clave para la supervivencia empresarial. Muchas empresas tuvieron que reevaluar rápidamente sus modelos de negocio, adoptando nuevas tecnologías y estrategias para mantenerse a flote. Por ejemplo, la digitalización, que incluye la adopción de comercio electrónico, teletrabajo y soluciones de tecnología de la información, se aceleró significativamente durante la pandemia. Como señala Martínez (2021), esta adaptación tecnológica no solo fue crucial para la continuidad de las operaciones durante los cierres, sino que también está configurando el futuro del entorno empresarial.

El apoyo gubernamental también jugó un papel vital en la mitigación del impacto económico de la pandemia. En España, el gobierno implementó una serie de medidas para apoyar a las empresas, incluyendo moratorias fiscales, líneas de crédito y subsidios para trabajadores afectados por ERTes (Expedientes de Regulación Temporal de Empleo). Estas medidas, como indica López-Cózar (2020), fueron esenciales para proporcionar un alivio financiero a las empresas y evitar una ola masiva de quiebras y despidos.

En definitiva, la pandemia de COVID-19 ha sido un catalizador de cambio para las empresas en España, resaltando la importancia de la innovación, la adaptabilidad y el apoyo gubernamental. Aunque el camino hacia la recuperación puede ser largo y desafiante, las lecciones aprendidas durante esta crisis serán fundamentales para el futuro del tejido empresarial español.

## OBJETIVOS Y EVOLUCIÓN

En este capítulo, se discutirá cómo los objetivos originales relacionados con el impacto de la pandemia en el ecosistema empresarial español se transformaron hacia un enfoque más amplio y detallado en el análisis de cómo diversas variables macroeconómicas influyen en la constitución y disolución de empresas. Además, se detallará cómo se formularon los nuevos objetivos y preguntas de investigación, subrayando el proceso de adaptación y reajuste metodológico que ha guiado el desarrollo y la dirección del análisis de datos en esta fase crucial del estudio.

### **Evolución**

La trayectoria del Trabajo de Fin de Grado (TFG) ha experimentado una notable evolución, reflejando un proceso de adaptación y refinamiento continuos que han estado

vinculados con los desafíos y las oportunidades descubiertas a lo largo del camino investigativo. Este proceso se ha caracterizado por un cambio significativo en el enfoque del estudio, motivado por las circunstancias emergentes y los hallazgos preliminares durante las etapas iniciales de la investigación.

Inicialmente, el objetivo del TFG se centraba en explorar el impacto de la pandemia de COVID-19 en el ecosistema empresarial español. Las preguntas de investigación originales buscaban comprender cómo se había desarrollado este ecosistema antes de la pandemia, así como identificar los efectos de la crisis sanitaria en la tasa de creación de empresas y en la resiliencia empresarial. Sin embargo, esta fase inicial, denominada **Fase 1: Búsqueda**, reveló una significativa limitación en términos de disponibilidad y profundidad de los datos específicos relacionados con el COVID-19, lo que impedía un análisis exhaustivo y riguroso.

Ante estos obstáculos, el proyecto dio un giro metodológico en la **Fase 2: Análisis**, donde se amplió el marco de estudio para incluir una variedad de variables macroeconómicas que pudieran influir en el ecosistema empresarial más allá del contexto pandémico. Este ajuste permitió aprovechar un espectro más amplio de datos disponibles, facilitando así una exploración más rica y diversificada de los factores que moldean el panorama empresarial español.

La **Fase 3: Especificación**, marcó una etapa decisiva en la evolución del TFG, donde se definieron con precisión las variables de estudio, enfocándose específicamente en las constituciones y disoluciones empresariales como indicadores clave del comportamiento del ecosistema empresarial. Este enfoque más focalizado permitió una investigación más detallada y especializada, orientada a descubrir las interacciones y dependencias entre las variables macroeconómicas y las tendencias empresariales.

A lo largo de estas fases, el TFG no solo transformó su dirección y alcance en respuesta a las limitaciones encontradas, sino que también se adaptó a los métodos de análisis más apropiados para abordar la nueva orientación del estudio. La adaptabilidad y la evolución del enfoque del TFG destacan la importancia de la flexibilidad en la investigación académica, especialmente en campos tan dinámicos y condicionados por factores externos como lo es el análisis empresarial y económico.

Este proceso evolutivo del TFG ha demostrado ser una parte integral del aprendizaje y el desarrollo académico del investigador, proporcionando una base sólida para la formulación de conclusiones significativas y recomendaciones bien fundamentadas, que serán detalladas en las secciones subsiguientes de este trabajo.

## Objetivos Planteados

Después de enfrentar desafíos iniciales y limitaciones de datos en las primeras etapas del TFG, se recalibraron los objetivos de investigación para adoptar un enfoque más robusto centrado en el análisis de variables macroeconómicas. Esta reorientación no solo mantiene la continuidad con las preocupaciones originales sobre la pandemia, sino que también amplía el espectro de estudio para incluir una comprensión más integral del ecosistema empresarial español bajo la influencia de diversas variables y crisis económicas. Originalmente, el proyecto intentó explorar el desarrollo del ecosistema empresarial antes y después de la pandemia de COVID-19, pero la falta de datos pertinentes llevó a modificar el enfoque hacia un análisis más amplio basado en variables macroeconómicas, definiendo nuevas preguntas de investigación más adecuadas al contexto y datos disponibles:

- *¿Cómo se ha desarrollado el ecosistema empresarial español en función del PIB e IPC, y qué implicaciones tienen estos indicadores en la estructura empresarial del país?*
- *¿Han afectado las variables macroeconómicas a la tasa de creación y disolución de empresas en España?*
- *¿Qué variables específicas influyen en la constitución y disolución de empresas?*
- *¿Cómo han respondido las empresas españolas ante las incertidumbres generadas por crisis macroeconómicas y cuáles son los factores clave para su resiliencia?*

Este TFG combina un análisis riguroso de datos con perspectivas obtenidas de entrevistas a expertos y una revisión exhaustiva de literatura académica y fuentes digitales confiables. El enfoque multidimensional no solo busca responder a las preguntas clave del estudio, sino que también pretende ofrecer una visión integral de cómo los factores económicos y las políticas gubernamentales interactúan para influir en el panorama empresarial del país, teniendo como objetivo proporcionar recomendaciones prácticas y bien fundamentadas que beneficien a emprendedores, empresarios y responsables de la formulación de políticas, enfocadas en fortalecer la resiliencia y promover el crecimiento sostenible del ecosistema empresarial en tiempos de crisis.

# PRIMERA PARTE: INGENIERÍA DEL DATO

## 1. Introducción a la ingeniería del dato

Esta parte del documento detalla el proceso de Extracción, Transformación y Carga (ETL) realizado para evaluar el impacto de COVID-19 en el tejido empresarial español. Se enfoca en cómo la pandemia y las políticas gubernamentales afectaron la creación y desarrollo de empresas. Los datos se limpiaron, estructuraron y analizaron meticulosamente, enfocándose en la precisión y corrección de discrepancias. Se seleccionaron variables y períodos específicos para alinearlos con los objetivos del Trabajo de Fin de Grado, empleando técnicas estadísticas para resumir y visualizar las tendencias. El capítulo concluye con un resumen que integra los hallazgos y describe el almacenamiento de datos post-ETL para futuros análisis.

## 2. Origen de los Datos

Los datos utilizados en este estudio provienen de fuentes confiables como el Instituto Nacional de Estadística (INE) y epdata, obtenidos mediante descarga directa para garantizar su integridad. El INE ofrece datos estadísticos que reflejan los aspectos socioeconómicos y demográficos de España, mientras que epdata complementa estos datos con visualizaciones y análisis detallados. Este método de recopilación mantiene los datos en su formato original, asegurando la transparencia y la reproducibilidad del proceso, elementos fundamentales en la investigación académica. Ambas fuentes son reconocidas por su rigor en la gestión de datos, proporcionando una base sólida y fiable para el análisis en este Trabajo de Fin de Grado.

## 3. Descripción de los Datos

### 3.1. Número total de observaciones y variables en los datasets (limpias)

En el marco de esta primera parte del Trabajo de Fin de Grado, se ha llevado a cabo un exhaustivo proceso de análisis y limpieza de datos, partiendo de un conjunto inicial de nueve bases de datos, denominadas "BBDD X", donde "X" representa el número asignado a cada una. A lo largo de este proceso, algunas bases de datos, específicamente "BBDD 3" y "BBDD 9", fueron descartadas debido a criterios de relevancia y calidad de los datos. Las bases de datos restantes, tras ser sometidas a un riguroso proceso de limpieza y transformación, han proporcionado un conjunto de datos depurado y estructurado, listo para

el análisis posterior. A continuación, se presenta una pequeña descripción detallada de las bases de datos limpias, ahora denominadas "LIMP.BBDD X", que se desarrollará de manera más extensa en el *Anexo 1*:

**Tabla 1.** Variables de LIMP.BBDD 1 clasificadas por tipo de variable e información sobre ellas

Nombre	Tipo de Dato	Descripción	Información Adicional
Column1	Integer	Número de observación.	Desde 0 hasta 9797.
Comunidad Autónoma	String	Nombre de las Comunidades Autónomas de España	Incluye el código numérico de la comunidad.
Forma jurídica	String	Tipo de forma jurídica de las sociedades.	Valores: "S.A.", "S.L.", "S. COM.", "S. COM. P.A. y S.C."
Sociedades o Capital desembolsado	String	Indica si hace referencia al número de sociedades o al capital desembolsado.	Valores: "Número de Sociedades", "capital desembolsado".
Año	Integer	Año de la observación, desde 2000 hasta 2022.	
Número Sociedades/Capital	Integer	Número de sociedades o capital.	Toma valor en función de la variable anterior
Población	Integer	Población correspondiente.	Según el año y la CCAA.

NOTA: "S.A.": Sociedad Anónima, "S.L.": Sociedad de Responsabilidad Limitada, "S. COM.", "S. COM. P.A. y S.C.": Sociedad Comanditaria, Sociedad Comanditaria por Acciones y Sociedad Colectiva

**Tabla 2.** Variables de LIMP.BBDD 2 clasificadas por tipo de variable e información sobre ellas

Nombre	Tipo de Dato	Descripción	Información Adicional
Column1	Integer	De observación 0 a la 242.	Índice de las filas.
Estados Sociedades	String	Estado de las sociedades.	"Constituidas" o "Disueltas".
Actividad económica	String	Actividad económica según la CNAE-2009.	
Año	Integer	Desde 2012 hasta 2022.	
Número Sociedades	Integer	Número de sociedades	Según estado, actividad y año.

NOTA: El CNAE-2009 (Clasificación Nacional de Actividades Económicas) es una clasificación que agrupa actividades económicas para fines estadísticos y administrativos.

### 3.2. Formato de los datos en bruto y cualquier transformación aplicada.

El proceso de transformación y limpieza de datos es un paso crucial en cualquier análisis estadístico. A continuación, se detalla cómo se han transformado los datos en bruto de varias bases de datos a formatos más estructurados y analíticamente útiles, describiendo las operaciones realizadas y añadiendo capturas de pantalla del código utilizado.

### **3.2.1) Base Datos 1:**

La base de datos "BBDD 1" contenía inicialmente 19.872 registros e incluía variables clave para el análisis como "Total Nacional", "Comunidades y Ciudades Autónomas", "Provincias", "Forma Jurídica", entre otras. La limpieza de esta base de datos fue un proceso meticuloso destinado a optimizar la utilidad de los datos para análisis más precisos. Se eliminaron columnas como "Total Nacional", que no eran relevantes para el enfoque del estudio. También se descartaron filas vacías en las columnas "Comunidades y Ciudades Autónomas" y "Provincias", así como aquellas filas con el valor "Total" en "Forma Jurídica", que no aportaban al análisis específico planeado.

Además, se eliminaron las entradas que contenían valores de "capital (en miles de euros) suscrito" en la variable "Número de sociedades y capital (en miles de euros)", para centrarse en datos más pertinentes y directamente relacionados con el objetivo del estudio. Posteriormente, se enriqueció la base de datos con datos adicionales para permitir comparaciones homogéneas. Se integraron datos poblacionales para cada comunidad y ciudad autónoma por cada periodo, utilizando dos bases de datos complementarias que abarcaban los años desde 2000 hasta 2022. Esto resultó en una base de datos consolidada y enriquecida que incluyó las variables esenciales de "Comunidades y Ciudades Autónomas", "Periodo" y "Población", preparada para análisis futuros más complejos y detallados.

### **3.2.2) Base Datos 2:**

La base de datos "BBDD 2" constaba de 3.102 registros y centraba su información en variables esenciales como "Estados Sociedades", "Actividad Económica", "Periodo" y "Total". La transformación de esta base de datos tuvo como objetivo su adaptación para análisis específicos, mejorando la claridad y aplicabilidad de la información. Se realizaron ajustes significativos, incluyendo la simplificación de la variable "Estados Sociedades" a "Constituidas", la modificación de denominaciones dentro de "Actividad Económica" para precisar las categorías económicas, y la reorganización de los datos por año, excluyendo los del año 2023 para enfocar el análisis en un rango temporal más pertinente. Adicionalmente, se implementó una función para calcular la suma total de sociedades por año, lo que facilitó una comprensión agregada de los datos. Estas acciones resultaron en la creación de "LIMP.BBDD 2", una base estructurada y optimizada para el análisis subsiguiente, proporcionando un recurso valioso para evaluar la evolución de las sociedades y actividades económicas a lo largo del tiempo.

### **3.2.3) Base Datos 4:**

La base de datos "BBDD 4", con 73,140 registros, incluía variables como "Clase de Disolución", "Provincias", "Periodo", y "Total". Para mejorar su utilidad en el análisis, se ajustaron los datos mediante varios pasos clave: se unieron DataFrames para asignar comunidades autónomas a provincias, facilitando análisis regionales más detallados; se reorganizaron los periodos mensuales por años en orden descendente para clarificar tendencias; se renombraron las columnas "Periodo" a "Año" y "Total" a "Número Sociedades" para reflejar mejor los datos anuales y describir explícitamente el total de sociedades disueltas. Estas transformaciones apuntaron a adaptar la base de datos a necesidades analíticas específicas, aumentando su claridad y usabilidad.

### **3.2.4) Base Datos 5:**

La quinta base de datos, "BBDD 5", originalmente compuesta por 2,376 registros, se enfocaba en variables como "Total Nacional", "Comunidades y Ciudades Autónomas", "Principales Indicadores", "Periodo" y "Total". Para ajustarla a los requisitos específicos del análisis y aumentar su funcionalidad, se implementaron varias modificaciones: se eliminó la columna "Total Nacional" por ser irrelevante; se renombró y limpió "Comunidades y Ciudades Autónomas" para garantizar datos completos; se filtraron y renombraron los "Principales Indicadores" para enfocarse únicamente en el "Índice de confianza empresarial armonizado" (ICEA), simplificando la variable a "Indicador"; se transformó "Periodo" de cuatrimestres a años completos para una visión temporal más clara; y se cambió el nombre de "Total" a "Valor Indicador", calculando su media anual por comunidad y año, y redondeando los valores para simplificar su presentación.

### **3.2.5) Base Datos 6 y 7:**

Las bases de datos 6 y 7, combinadas debido a sus estructuras semi-estructuradas y contenido similar, fueron meticulosamente procesadas para asegurar la utilidad y relevancia de los datos para el análisis. La información contenida en formatos de Excel dispersos requería un enfoque detallado para su organización y limpieza.

Inicialmente, se seleccionaron variables críticas para el estudio del impacto de los ERTes, enfocándose en categorías como actividad económica, fechas, comunidades autónomas y detalles de los ERTes. Este proceso de selección se orientó a maximizar la relevancia de los datos para los objetivos analíticos del estudio, centrados principalmente en los años 2021 y 2022.



La manipulación de datos semi-estructurados implicó técnicas manuales utilizando funciones de Excel para consolidar y estructurar la información de manera coherente. Se extrajeron de múltiples hojas y tablas dentro de los archivos.

Finalmente, se creó una tabla unificada que integraba todas las variables seleccionadas de ambas bases de datos, ofreciendo una visión de la situación de los ERTes en España durante el período estudiado.

### **3.2.6) Base Datos 8:**

La base de datos 8 se diseñó para facilitar el análisis de la afiliación a la seguridad social en diferentes periodos y constaba de 670 registros con variables clave como "Año", "Periodo" y "Afiliados a la Seguridad Social". Para mejorar la usabilidad de los datos, se realizaron modificaciones significativas en la estructura y el formato. La variable "Periodo", que inicialmente presentaba fechas en un formato no óptimo, se ajustó a un formato más estándar y manejable, 'dd/mm', facilitando la interpretación y el manejo de la información temporal. Además, para simplificar el análisis y mejorar la coherencia de los datos, se combinaron las variables "Periodo" y "Año" en una nueva variable denominada "Fecha". Este cambio consolidó los detalles temporales en un campo único, lo que simplificó el análisis posterior. Finalmente, se optó por reducir el conjunto de datos a dos variables fundamentales: "Fecha" y "Afiliados a la Seguridad Social", eliminando cualquier información redundante o no esencial. Este enfoque enfocado permitió un análisis más claro de la evolución de la afiliación a la seguridad social, proporcionando una base sólida para investigar tendencias laborales y socioeconómicas significativas.

## **3.4. Periodicidad y temporalidad de los datos.**

La consideración de la periodicidad y la temporalidad de los datos ha sido un aspecto crucial en este análisis. Las cinco primeras bases de datos (BBDD 1 a BBDD 5) están organizadas anualmente, lo que permite un análisis eficaz de tendencias a largo plazo y facilita comparaciones año a año. La BBDD 1, que abarca desde el 2000 hasta el 2022, ofrece una perspectiva extensa de más de dos décadas, útil para evaluar el impacto prolongado de la pandemia. Las BBDD 2 y 3, que cubren desde 2012 hasta 2022, y las BBDD 4 y 5, desde 2013 hasta 2022, proporcionan datos fundamentales para analizar las dinámicas empresariales recientes, incluyendo las adaptaciones a la crisis sanitaria. Contrariamente, la BBDD 6, con fechas detalladas desde el "01/03/2020" hasta el "30/12/2021", permite un enfoque más minucioso en los efectos inmediatos de la COVID-19, capturando las respuestas a corto plazo a las restricciones y políticas de salud implementadas durante la pandemia.

## 4. Justificación de la Elección de Variables

### 4.1. Explicación de la selección de variables.

Las variables seleccionadas incluyen "Año", "Número de Sociedades", "Indicador", "Tipo de ERTE", "Número de ERTES", "Comunidad Autónoma", entre otras. Estas variables fueron escogidas por su relevancia directa en la evaluación del entorno empresarial y su evolución durante y después de la pandemia. Por ejemplo:

*Año:* Permite realizar comparaciones temporales y evaluar tendencias antes, durante y después de la pandemia.

*Número de Sociedades:* Es fundamental para medir la tasa de creación y disolución de empresas, proporcionando una visión clara de cómo ha fluctuado el tejido empresarial.

*Indicador (ICEA):* Ofrece una perspectiva sobre la confianza empresarial, lo cual es un termómetro del clima económico y empresarial.

*Tipo de ERTE y Número de ERTES:* Estas variables son cruciales para evaluar el impacto de las medidas gubernamentales, especialmente en lo que respecta a la viabilidad y continuidad de las empresas durante la pandemia.

*Comunidad Autónoma:* Permite realizar análisis regionales entendiendo mediante diferentes áreas geográficas como han sido afectadas y han respondido a la crisis las empresas.

### 4.2. Relación con los objetivos del TFG.

Para evaluar la dinámica empresarial durante la pandemia, se utilizan variables como "Número de Sociedades" y "Año" para determinar las tasas de creación y disolución de empresas, revelando el impacto de la crisis en la iniciativa empresarial española. Asimismo, mediante la variable "Actividad Económica", se analiza la respuesta de distintos sectores a la pandemia, identificando aquellos que han resistido mejor o sufrido más. La efectividad de las políticas gubernamentales, como los ERTES, se estudia a través de variables como "Tipo de ERTE" y "Número de ERTES", para comprender su papel en la supervivencia empresarial durante los periodos de restricciones. Además, el análisis del "Indicador" (ICEA) junto con el "Número de Sociedades" a lo largo del tiempo permite evaluar la resiliencia y la capacidad de adaptación del tejido empresarial ante crisis.

## 5. Análisis Exploratorio de Datos

### 5.1. Descripción de medidas de tendencia central y dispersión para las variables cuantitativas.

#### 5.1.2 Medidas de tendencia central

En el marco del Trabajo de Fin de Grado, se procederá a realizar un análisis exhaustivo de las variables cuantitativas seleccionadas mediante la aplicación de medidas de tendencia central. Este análisis tiene como objetivo principal proporcionar una comprensión detallada de la distribución central de los datos, lo cual es esencial para identificar patrones, tendencias y posibles anomalías dentro del conjunto de datos. Para cada una de estas variables, se calcularán las siguientes medidas de tendencia central:

- *Media*: Esta medida proporcionará el promedio de los valores para cada variable, ofreciendo una visión general del valor central en torno al cual se distribuyen los datos.
- *Mediana*: Al identificar el valor medio en el conjunto de datos ordenado, la mediana nos permitirá entender el punto central de la distribución, minimizando el efecto de valores atípicos extremos.
- *Moda*: La identificación de los valores más frecuentes en el conjunto de datos nos ayudará a comprender las tendencias predominantes y las preferencias dentro del tejido empresarial.

**Tabla 3.** Medidas de tendencia central de la primera base de datos

Nombre	Media	Mediana	Moda
Año	2011	2011	2000
Número Sociedades/Capital	90,53	6	0
Población	2397012,45	1464847	1107220

*Para un análisis detallado de las medidas de tendencia central, véase el Anexo 1.*

#### 5.1.3 Medidas de dispersión

A continuación, se abordará una dimensión complementaria al estudio de las variables cuantitativas mediante la aplicación de medidas de dispersión. Este enfoque se centra en evaluar la variabilidad o dispersión de los datos alrededor de un valor central, lo cual es crucial para comprender el grado de variación dentro del conjunto de datos y, por

ende, la consistencia o heterogeneidad de las observaciones. Las medidas de dispersión seleccionadas para este análisis son:

- *Rango*: Esta medida refleja la diferencia entre el valor máximo y mínimo dentro del conjunto de datos para cada variable. El rango proporciona una visión inicial de la amplitud de la variabilidad, aunque es sensible a valores extremos.
- *Varianza*: Cuantifica la variabilidad de los datos calculando el promedio de los cuadrados de las desviaciones respecto a la media. Aunque proporciona una medida precisa de la dispersión, su interpretación puede ser menos intuitiva debido a que las unidades están al cuadrado respecto a las de la variable original.
- *Desviación Estándar*: Representa la raíz cuadrada de la varianza y mide la dispersión de los datos respecto a su media. Una desviación estándar baja indica que los datos tienden a estar cerca de la media, mientras que una desviación estándar alta señala una mayor dispersión alrededor de la media.
- *Cuartiles 1 y 3*: Los cuartiles dividen el conjunto de datos ordenado en cuatro partes iguales. El primer cuartil (Q1), valor por debajo del cual se encuentra el 25% de los datos y el tercer cuartil (Q3), el valor por debajo del cual se sitúa el 75% de los datos, son particularmente significativos en el análisis estadístico. Estos puntos de corte proporcionan una visión clara de la distribución de los datos, permitiendo identificar dónde se concentran la mayoría de las observaciones y cómo se distribuyen los valores tanto en la parte inferior como en la superior de la muestra.
- *Rango Intercuartílico (IQR)*: El rango intercuartílico se define como la diferencia entre el tercer cuartil (Q3) y el primer cuartil (Q1). Esta medida de dispersión es especialmente útil para evaluar la variabilidad de los datos minimizando el impacto de los valores atípicos o extremos.

**Tabla 4.** Medidas de dispersión de la primera base de datos.

Nombre	Rango	Varianza	Desviación Estándar	Rango Intercuartílico
Año	22	44,02	6,63	12
Número Sociedades/Capital	997	38209,70	195,47	66,74
Población	8445679	5693262396225,57	2386055,82	2063773

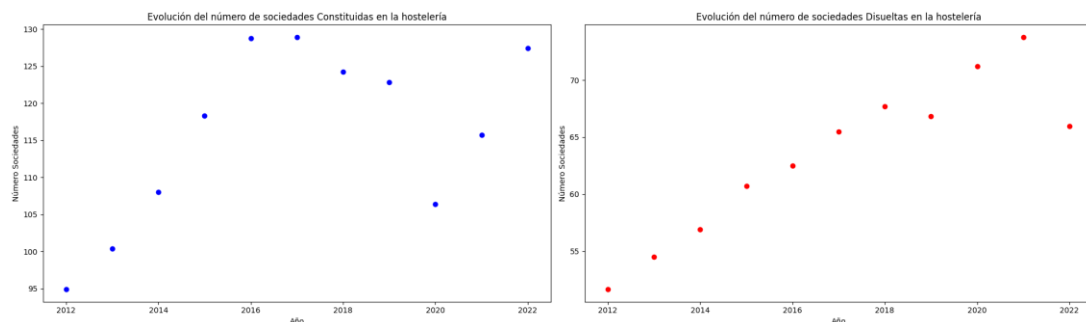
Para un análisis detallado de las medidas de dispersión, véase el Anexo 1.

## 6. Gráficos Descriptivos

### 6.1. Gráficos sobre variables y relaciones.

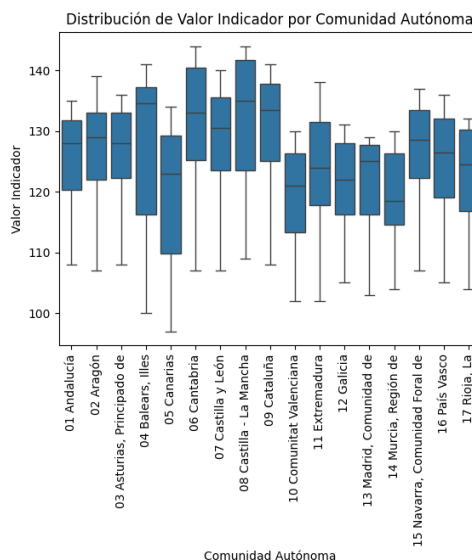
En esta investigación, se emplea un enfoque meticuloso para analizar el sector empresarial español. Se utilizan gráficos descriptivos como herramientas clave para una interpretación efectiva y directa de la complejidad de los datos. Se seleccionaron varios tipos de representaciones visuales, incluyendo diagramas de dispersión, boxplots y histogramas, que son esenciales para detallar las características y tendencias tanto de variables cuantitativas como categóricas. Estos gráficos son fundamentales para profundizar en la comprensión de las dinámicas empresariales durante la crisis sanitaria global.

*Diagrama de Dispersión:* El diagrama de dispersión se utiliza para examinar la relación entre dos variables cuantitativas, como el año y el número de sociedades constituidas en relación con variables específicas como la comunidad autónoma o la forma jurídica. Este gráfico es esencial para analizar casos concretos y evitar la sobrecarga de información, centrando el análisis en tipos específicos de empresas para aumentar su valor analítico.



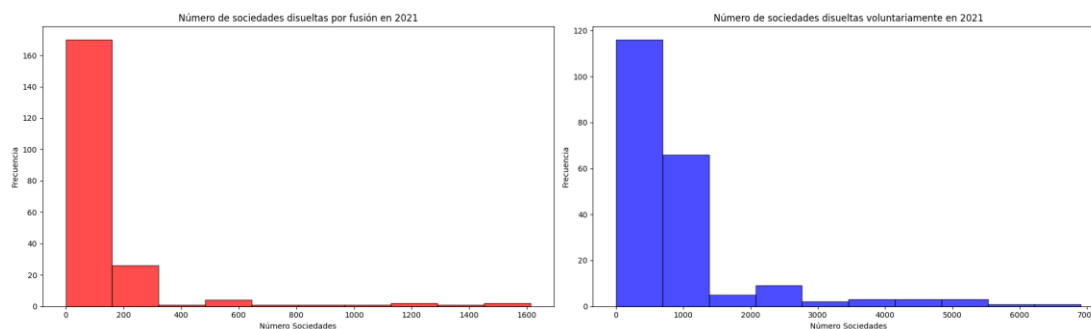
**Ilustración 1.** Captura de pantalla de los diagramas de dispersión correspondientes a las evoluciones de las sociedades Constituidas y disueltas en el sector de la hostelería desde el año 2012 al 2022.

*Diagrama de Cajas (Boxplot):* Este tipo de gráfico proporciona una visión comprensiva de la distribución de los datos, resaltando aspectos clave como la mediana, los cuartiles y los outliers. Se empleará para analizar la variabilidad de variables como el valor del indicador ICEA o el capital desembolsado a lo largo de diferentes comunidades autónomas o sectores económicos. La selección cuidadosa de variables para este análisis es esencial para prevenir interpretaciones erróneas y asegurar que la información presentada sea pertinente y accesible.



**Ilustración 2.** Captura de pantalla del diagrama de cajas correspondiente a la distribución del valor indicador “ICEA” según la comunidad autónoma

**Histograma:** A través de los histogramas, se visualiza la distribución de frecuencias de una variable cuantitativa, lo que permite identificar patrones como la distribución normal o sesgos en los datos. Al igual que con los diagramas de dispersión, se focalizará el análisis en aspectos concretos del conjunto de datos, tales como el número de sociedades disueltas de manera voluntaria en 2021. Este enfoque dirigido garantiza que los histogramas proporcionen insights claros y específicos sobre las dinámicas empresariales bajo estudio.



**Ilustración 3.** Captura de pantalla del histograma correspondiente a la evolución del número de sociedades disueltas voluntariamente en 2021 en todas las comunidades autónomas españolas.

## 6.2. Patrones, tendencias o correlaciones observadas en los gráficos.

Reflejado en la *Ilustración 1* del TFG, se presenta un análisis comparativo de la evolución de las sociedades constituidas y disueltas en el sector de la hostelería en España, desde el año 2012 hasta 2022. La elección de este sector es particularmente pertinente dado que la hostelería ha sido uno de los sectores más visiblemente afectados por la pandemia del COVID-19, y se busca comprender las implicaciones que dicho evento ha tenido en la dinámica empresarial.

En el gráfico de la izquierda, que muestra las sociedades constituidas con puntos azules, se observa un crecimiento exponencial hasta 2017, seguido de un descenso gradual hasta 2019. Sin embargo, entre 2019 y 2020, se produce una caída drástica en el número de nuevas sociedades, lo que se alinea temporalmente con el inicio de la pandemia y las consecuentes restricciones operativas impuestas al sector. Aunque en 2021 se aprecia una recuperación, esta no alcanza los niveles previos al 2019, y en 2022, sin embargo, a partir del 2022 los valores parecen indicar un retorno a la normalidad.

El gráfico de la derecha, representado con puntos rojos, refleja una tendencia creciente en la disolución de sociedades desde 2012 hasta 2021. Es particularmente interesante observar que durante los años 2019 y 2022 no se aprecia un incremento, lo cual podría sugerir una estabilización temporal en la desaparición de empresas. El estancamiento en 2019 podría interpretarse como una consolidación previa a la crisis sanitaria. En contraste, el año 2022, que no muestra un aumento, podría estar indicando el inicio de una recuperación del sector o el resultado de las medidas de apoyo gubernamentales que han permitido a las empresas resistir las adversidades.

La *Ilustración 2* muestra la distribución del Índice de Confianza Empresarial Armonizado (ICEA) por comunidades autónomas en España mediante un diagrama de cajas. Este índice mide la percepción del sector empresarial sobre el clima económico, revelando tendencias regionales y variaciones en la confianza empresarial. Los boxplots ilustran la mediana, cuartiles y la variabilidad del ICEA en cada región. Se observa una mayor variabilidad en comunidades como Baleares y Canarias, mientras que en Asturias y Extremadura la percepción es más uniforme.

Por último, la *Ilustración 3* expone dos histogramas que representan distintas modalidades de disolución de sociedades en España durante el año 2021, distinguiendo entre disoluciones por fusión y disoluciones voluntarias.

El primer histograma, en tono rojo, muestra la frecuencia de sociedades disueltas por fusión, concentrándose principalmente entre 0 y 300 casos. Esto sugiere que las fusiones frecuentemente involucran a un número menor de sociedades, indicando que las grandes fusiones son menos comunes.

El segundo histograma, en azul, representa la frecuencia de disoluciones voluntarias de sociedades, destacando una acumulación en el rango de 0 a 1500, con una disminución posterior. Esto señala que las disoluciones voluntarias superan en número a las fusiones, reflejando una preferencia por decisiones autónomas en el contexto económico actual, posiblemente influenciado por la pandemia de COVID-19.

# SEGUNDA PARTE: ANÁLISIS DEL DATO

## 8. Introducción al análisis del dato

Este apartado describe la aplicación de análisis de datos avanzados para entender las dinámicas que afectan las disoluciones empresariales en España, destacando la importancia de estos métodos en la investigación y toma de decisiones. Se detalla cómo los datos son seleccionados y preparados para análisis mediante modelos que buscan patrones y causas, empleando un enfoque multidimensional que abarca variables económicas, sociales y de mercado. Se evalúa la precisión de los modelos con métodos estadísticos y se utilizan visualizaciones para clarificar los resultados. El análisis concluye con recomendaciones prácticas basadas en los hallazgos, buscando contribuir al conocimiento académico y proporcionar herramientas útiles para la práctica empresarial y el desarrollo de un entorno económico resiliente.

## 9. Selección y Preparación de Datos para el Análisis

La selección y preparación de datos constituyen etapas cruciales en el proceso de análisis de datos, especialmente cuando se abordan cuestiones complejas como las situaciones que influyen en las disoluciones empresariales. Este trabajo se ha fundamentado en el análisis exhaustivo de tres bases de datos principales, cada una de ellas derivada y refinada a partir de conjuntos de datos más amplios, con el objetivo de explorar distintas facetas del fenómeno en estudio.

La primera base de datos, inicialmente denominada "dfbbdd1" y posteriormente segmentada en "numsoc" y "capdes", se centró en recopilar información relativa al número de sociedades creadas por tipo de sociedad, año y comunidad autónoma, así como el capital desembolsado por estas sociedades, complementado con datos demográficos por comunidad autónoma. Esta división permitió abordar dos análisis de regresión lineal simple, orientados a evaluar el impacto de la población en la creación de empresas y en el capital desembolsado, respectivamente, proporcionando una visión detallada de cómo la demografía puede influir en el tejido empresarial.

En el desarrollo de la investigación para este Trabajo de Fin de Grado, se configuraron dos bases de datos fundamentales para el análisis de las dinámicas empresariales en España. La primera, denominada "disolución", compila el número de empresas disueltas por tipo de disolución en cada comunidad autónoma desde el año 2012



hasta el 2022. Complementariamente, se creó una segunda base de datos, llamada "macrospain", que recoge indicadores económicos esenciales como el Producto Interno Bruto (PIB) y el Índice de Precios al Consumidor (IPC) por comunidad autónoma y año. El objetivo de vincular estos conjuntos de datos era llevar a cabo una regresión lineal múltiple para investigar cómo las condiciones económicas impactan en la tasa de disolución empresarial.

De manera paralela y con la misma metodología, se desarrolló otra base de datos, "regmul2", destinada a registrar el número de empresas constituidas en cada comunidad autónoma durante el mismo periodo. Esta base de datos se orienta a explorar la relación entre los mismos indicadores económicos y la tasa de creación de empresas, permitiendo un análisis comparativo entre los factores que influyen tanto en la constitución como en la disolución de empresas dentro del territorio español. Esta dualidad de bases de datos enriquece significativamente el estudio, permitiendo una comprensión más completa de las fuerzas económicas que moldean el panorama empresarial del país.

Para profundizar en el análisis de las dinámicas empresariales en España, se creó una tercera base de datos denominada "dismac", destinada a ampliar el espectro de variables macroeconómicas analizadas en relación con las disoluciones empresariales. Esta base de datos incluye indicadores tales como la deuda pública, el déficit público, el gasto público, los ingresos fiscales, el turismo y las reservas nacionales, junto con el Índice de Confianza Empresarial Armonizado (ICEA), para cada comunidad autónoma y año. Este enfoque multidimensional ofrece una visión más rica y detallada de cómo diversos factores económicos y sociales pueden influir en la estabilidad y continuidad de las empresas en España.

De manera similar, se creó una base de datos complementaria llamada "consmac", que aplica la misma metodología, pero se centra en las constituciones empresariales. "consmac" también integra las mismas variables macroeconómicas y busca explorar cómo estos factores influyen en la creación de nuevas empresas, permitiendo un análisis comparativo entre los impulsores de las constituciones y disoluciones empresariales. Este enfoque paralelo facilita una comprensión integral de los efectos de las políticas económicas y las condiciones de mercado en el tejido empresarial del país.

Por último, para complementar y enriquecer aún más el análisis sobre las dinámicas empresariales en España, se han incorporado dos nuevas bases de datos: "consext" y "disext", correspondientes a las constituciones y disoluciones empresariales, respectivamente. Estas bases de datos han sido diseñadas para explorar la relación entre un amplio espectro de variables y las dinámicas de constitución y disolución de empresas.

Incluyen variables tanto económicas como sociales, ampliando considerablemente el rango de factores considerados en el análisis. Entre estas variables se encuentran la cotización del IBEX, del EUR y del NASDAQ, que reflejan el comportamiento del mercado financiero; indicadores macroeconómicos como el déficit público, el gasto público y los ingresos fiscales; y variables demográficas como nacimientos, matrimonios, defunciones y población total.

La inclusión de estas variables busca proporcionar una visión más holística y detallada de los factores que pueden influir en la creación y cierre de empresas en España. Se parte del supuesto de que, a mayor número de variables analizadas, mayor será la capacidad del modelo para explicar las fluctuaciones en las tasas de constituciones y disoluciones empresariales. Con "consect" y "disect", se pretende no solo identificar las tendencias y patrones más evidentes, sino también descubrir conexiones menos obvias que puedan surgir de la interacción entre el entorno económico, el mercado financiero y la dinámica social, ofreciendo así una comprensión más profunda y matizada de los desafíos y oportunidades dentro del ecosistema empresarial español.

Cada una de estas bases de datos fue sometida a un riguroso proceso de selección y preparación, que incluyó la limpieza de datos, la gestión de valores faltantes, la transformación de variables y la verificación de la calidad de los datos. Este proceso aseguró que la información utilizada en los análisis fuera de la más alta calidad y relevancia, permitiendo así obtener resultados confiables. La meticulosa preparación de los datos subraya la importancia de una base sólida para cualquier análisis de datos, especialmente cuando se abordan cuestiones de complejidad y relevancia como las que conciernen a las disoluciones empresariales en el contexto económico y social de España.

## **10. Modelos Analíticos: Desarrollo y Aplicación (Regresiones Lineales)**

### **10.1. Modelo Analítico Supervisado (Regresiones Lineales Simples)**

#### **Explicación del Modelo:**

En el análisis del comportamiento empresarial y económico, la regresión lineal simple emerge como una herramienta analítica fundamental, especialmente cuando el objetivo es explorar la relación entre dos variables específicas. Este modelo supervisado se seleccionó con el propósito de investigar cómo la variable independiente, en este caso, la población de una comunidad autónoma puede influir en la variable dependiente, que para la primera

regresión se define como el número de sociedades creadas y para la segunda como el capital desembolsado en miles de euros.

La regresión lineal simple es un método estadístico fundamental utilizado para modelar y analizar las relaciones entre dos variables cuantitativas: una variable independiente (o explicativa) y una variable dependiente (o respuesta). Este modelo asume que existe una relación lineal entre estas variables, la cual puede ser descrita mediante una ecuación de la forma:

$$y = \beta_0 + \beta_1 x + \epsilon$$

Donde:

- $y$  representa la variable dependiente.
- $x$  representa la variable independiente.
- $\beta_0$  es el término de intercepción, que indica el valor de  $y$  cuando  $x$  es 0.
- $\beta_1$  es el coeficiente de la pendiente, que indica el cambio en  $y$  por cada unidad de cambio en  $x$ .
- $\epsilon$  es el término de error, que representa la variación en  $y$  que no puede ser explicada por la relación lineal con  $x$ .

El método de mínimos cuadrados es utilizado para estimar los parámetros  $\beta_0$  y  $\beta_1$  de la regresión. Este método busca minimizar la suma de los cuadrados de las diferencias (residuos) entre los valores observados de " $y$ " y los valores predichos por el modelo. Matemáticamente, se busca minimizar la función:

$$S(\beta_0, \beta_1) = \sum_{i=1}^n (y_i - (\beta_0 + \beta_1 x_i))^2$$

Donde  $y_i$  y  $x_i$  son los valores observados de la variable dependiente e independiente, respectivamente, y  $n$  es el número de observaciones. La solución a este problema de optimización nos da los estimadores de mínimos cuadrados para  $\beta_0$  y  $\beta_1$ , que son:

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

$$\hat{\beta}_0 = \bar{y} - \beta_1 \bar{x}$$

Donde  $\bar{x}$  e  $\bar{y}$  son las medidas de las variables independiente y dependiente, respectivamente.

El  $\beta_1$  coeficiente indica la pendiente de la línea de regresión y representa el cambio esperado en  $y$  por cada unidad de cambio de  $x$ . Un valor positivo de  $\beta_1$  indica una relación directa entre  $x$  e  $y$ , mientras que un valor negativo indica una relación inversa.

El término de intercepción  $\beta_0$  indica el valor esperado de  $y$  cuando  $x$  es 0. Este valor puede no tener siempre una interpretación práctica, especialmente si  $x = 0$  no se encuentra dentro del rango de los datos observados.

### **Proceso de Selección de Variables:**

En el análisis mediante regresión lineal simple, se seleccionaron las variables más adecuadas para reflejar la relación de interés. Se utilizó la población de las comunidades autónomas como variable independiente, considerando su influencia en la creación de empresas y en la capacidad económica regional. Como variables dependientes, se eligió el número de sociedades creadas para una regresión y el capital desembolsado en miles de euros para la otra. Para efectuar el análisis, la base de datos se dividió en dos subconjuntos enfocados en cada variable dependiente respectivamente, permitiendo un examen detallado de cada variable. Este proceso fue clave para capturar efectivamente la relación entre la demografía regional y la actividad económica empresarial, y facilitó la interpretación de los resultados y la formulación de conclusiones pertinentes.

### **Desarrollo de modelos:**

#### **Regresión 1: Número de Sociedades**

En el análisis inicial para explorar la relación entre la población de las comunidades autónomas y el número de sociedades creadas, se utilizaron gráficos de dispersión con líneas de tendencia suavizadas. Estos gráficos ayudaron a visualizar la distribución y detectar patrones, aunque sugirieron una relación potencialmente no lineal o débil debido a la dispersión y suavidad de la línea de tendencia.

Para profundizar en este análisis, se empleó la función `corr()` de pandas para calcular la correlación entre 'Población' y 'Número de Sociedades', complementada con el cálculo del coeficiente de correlación de Pearson y su p-valor asociado. Los resultados arrojaron una correlación de -0.008 y un p-valor de 0.761, indicando estadísticamente la ausencia de una relación lineal significativa entre ambas variables.

El ajuste de las variables a una distribución normal fue verificado a través de gráficos de densidad y pruebas de normalidad, incluyendo Shapiro-Wilk, Anderson-Darling y

D'Agostino's  $K^2$ , todos mostrando distribuciones asimétricas y p-valores bajos que rechazan la normalidad. A pesar de esto, se procedió a ajustar un modelo de regresión lineal simple, que reveló un R-cuadrado cercano a cero y coeficientes de regresión sin significancia estadística, confirmando la falta de una relación lineal significativa entre la población y el número de sociedades.

## **Regresión 2: Capital Desembolsado**

Para explorar la relación entre la población de las comunidades autónomas y el capital desembolsado en la creación de sociedades, se utilizaron gráficos de dispersión con líneas de tendencia suavizadas. Estos gráficos proporcionaron una vista preliminar que sugirió la existencia de una relación que requería análisis adicionales. Se calculó la correlación entre 'Población' y 'Capital' usando la función `corr()` de pandas, complementada con el coeficiente de correlación de Pearson, resultando en una correlación de 0.20 con un p-valor de  $4.639e-14$ , indicando una relación lineal positiva estadísticamente significativa, aunque de fuerza moderada.

El análisis de la normalidad de las distribuciones mediante gráficos de densidad y pruebas estadísticas mostró asimetría en ambas variables, con pruebas de Shapiro-Wilk, Anderson-Darling y D'Agostino's  $K^2$  confirmando la no normalidad con p-valores muy bajos. Esto subrayó la falta de ajuste a una distribución normal, crucial para algunas técnicas estadísticas, y sugirió la necesidad de enfoques más robustos para el análisis.

A pesar de estos retos, se ajustó un modelo de regresión lineal simple para investigar más a fondo la relación entre la población y el capital desembolsado. El modelo ajustado mostró un R-cuadrado de 0.043, indicando que solo un pequeño porcentaje de la variabilidad del capital desembolsado se explica por la población. Sin embargo, el coeficiente asociado a la población fue significativo, confirmando una conexión positiva, aunque moderada entre estas variables. Este hallazgo resalta la importancia de considerar la significancia estadística y la relevancia práctica de las relaciones en estudios económicos.

## **Evaluación del Modelo con Medidas de Error/Precisión Específicas:**

### **Regresión 1: Número de Sociedades**

El modelo analizado presenta un R-cuadrado de 0.000, lo que indica una capacidad explicativa nula respecto a la relación entre la población y el número de sociedades creadas. Este resultado se ve acentuado por un R-cuadrado ajustado de -0.001 y un F-statistic de

0.09227 con un p-valor de 0.761, reflejando que el modelo no aporta una mejora significativa sobre una predicción al azar y carece de significancia estadística.

La confiabilidad del modelo también es baja, evidenciada por un alto valor de condición de  $4.79e+06$  que sugiere una posible multicolinealidad, aunque esto es menos probable en modelos de regresión simple. Los coeficientes de regresión y sus intervalos de confianza indican que, mientras el intercepto es significativo, la pendiente para la población no lo es, mostrando un intervalo que incluye el cero y un p-valor elevado. Esto se acompaña de una alta variabilidad en los residuos, demostrando que el modelo falla en capturar la esencia de los datos.

En resumen, el modelo de regresión lineal simple examinado no es efectivo para predecir el número de sociedades a partir de la población, mostrando una falta de precisión y significancia estadística. Esto sugiere que variables adicionales y análisis más complejos son necesarios para comprender adecuadamente los factores que influyen en la creación de sociedades, subrayando la necesidad de enfoques más integrales en investigaciones futuras.

## **Regresión 2: Capital Desembolsado**

El modelo de regresión lineal muestra un R-cuadrado de 0.043, lo que implica que alrededor del 4.3% de la variabilidad del capital desembolsado se explica mediante la población, indicando una relación positiva, pero de magnitud limitada entre estas variables. El R-cuadrado ajustado es de 0.042, confirmando una ligera eficacia predictiva del modelo basado en la población. Además, un F-statistic de 58.15 con un p-valor muy bajo ( $4.64e-14$ ) señala la significancia estadística del modelo, aunque la fuerza de la relación es relativamente débil.

La confiabilidad del modelo se ve comprometida por un alto valor de condición de  $4.79e+06$ , sugiriendo posibles problemas de multicolinealidad o errores numéricos que podrían afectar la precisión de los coeficientes. A pesar de esto, tanto el intercepto como la pendiente relacionada con la población son estadísticamente significativos, indicando que incrementos en la población se correlacionan con aumentos en el capital desembolsado. Sin embargo, los residuos del modelo y la suma de cuadrados de estos muestran que hay una cantidad considerable de variabilidad no explicada.

En conclusión, el modelo sugiere una conexión positiva entre población y capital desembolsado, pero con una capacidad explicativa limitada, apuntando a que factores no incluidos en este análisis podrían influir significativamente en el capital desembolsado. La presencia de un alto valor de condición y la varianza limitada explicada por el modelo

subrayan la necesidad de profundizar más en el análisis e incluir variables adicionales para capturar de manera más completa las dinámicas que afectan el capital desembolsado.

## 10.2. Modelo Analítico Supervisado (Regresiones Lineales Múltiples)

### Explicación del Modelo:

La regresión lineal múltiple, al igual que su contraparte simple, es una herramienta estadística esencial en el análisis de datos, especialmente útil para explorar la relación entre una variable dependiente y múltiples variables independientes. Este modelo supervisado permite investigar cómo varias variables independientes, como la población y el PIB de una comunidad autónoma, pueden influir conjuntamente en una variable dependiente, como el número de empresas constituidas o el capital desembolsado en miles de euros.

El modelo de regresión lineal múltiple se expresa mediante la siguiente ecuación:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \epsilon$$

Donde:

- $y$  representa la variable dependiente.
- $x_1, x_2, \dots, x_k$  representan las variables independientes.
- $\beta_0$  es el término de intercepción, que indica el valor de  $y$  cuando las  $x$  son 0.
- $\beta_1, \beta_2, \dots, \beta_k$  son los coeficientes de las variables independientes, que miden el cambio en  $y$  asociado a una unidad de cambio en cada  $x$ .
- $\epsilon$  es el término de error, que capta toda la variabilidad en  $y$  que no es explicada por las variables independientes.

El método de los mínimos cuadrados ordinarios también se utiliza aquí para estimar los coeficientes, buscando minimizar la suma de los cuadrados de los residuos, es decir, las diferencias entre los valores observados y los valores predichos por el modelo. La función a minimizar es:

$$S(\beta_0, \beta_1, \dots, \beta_k) = \sum_{i=1}^n (y_i - (\beta_0 + \beta_1 x_{1i} + \dots + \beta_k x_{ki}))^2$$

La regresión lineal múltiple es potente por su capacidad de ajustar múltiples variables, permitiendo un análisis más detallado y una mejor comprensión de cómo diversos factores afectan la variable objetivo. Los coeficientes obtenidos ofrecen una medida de la influencia de cada variable independiente, ajustada por la presencia de otras en el modelo.

### 10.2.1 PRIMERA REGRESIÓN LINEAL MÚLTIPLE (IPC y PIB)

#### Proceso de Selección de Variables

La selección del Índice de Precios al Consumo (IPC) y el Producto Interno Bruto (PIB) como predictores se fundamentó en la suposición de que estos indicadores de la salud económica influyen significativamente tanto las disoluciones como las constituciones empresariales. El IPC, indicativo de la inflación, y el PIB, reflejo del desempeño económico global, son clave para comprender el ambiente en que operan las empresas. La elección se basó en un análisis exploratorio y revisiones de literatura que sugieren una posible correlación entre estos factores económicos y el comportamiento empresarial. Se teorizó que un PIB creciente podría incentivar la fundación de nuevas empresas, mientras que un IPC alto podría reprimir las inversiones por el incremento en los costos que implica la inflación. Esta interacción destaca la importancia de estudiar cómo cada indicador afecta no solo las tasas de cierre de empresas sino también el surgimiento de nuevas.

#### Desarrollo del Modelo

En la elaboración del modelo para estudiar las disoluciones y constituciones empresariales, se utilizó un enfoque preliminar de análisis gráfico para evaluar la relación lineal entre el IPC, el PIB y las dinámicas empresariales. A pesar de que los coeficientes de correlación fueron bajos y no estadísticamente significativos (IPC: -0.0076 a 0.0179, PIB: 0.0202 a 0.0273), se decidió incluir estas variables en un modelo de regresión múltiple para un análisis más comprensivo, esperando capturar efectos combinados y más complejos.

Durante la preparación del modelo, se confirmó la no normalidad de las variables mediante pruebas significativas como Shapiro-Wilk y Anderson-Darling, destacando fuertes evidencias de distribuciones no normales tanto para disoluciones como para constituciones empresariales. A pesar de esta falta de normalidad, se procedió con la regresión lineal múltiple, fundamentándose en la robustez del modelo frente a este supuesto, apoyada por la visualización de posibles relaciones lineales débiles.

Los modelos finales para las disoluciones y constituciones empresariales indicaron una relación débil entre las variables económicas y el comportamiento empresarial, con coeficientes muy bajos y no significativos para IPC y PIB. Esto sugiere que otros factores no considerados podrían estar influyendo significativamente, resaltando la necesidad de explorar más variables y contextos específicos para entender completamente cómo las condiciones económicas afectan la dinámica empresarial. Ambos modelos, a pesar de sus bajos R-cuadrados, proporcionan insights importantes para futuras investigaciones.



### **Evaluación del Modelo con Medidas de Error/Precisión Específicas:**

La evaluación de la calidad de los modelos de regresión para las disoluciones y constituciones empresariales en España reveló una capacidad explicativa muy limitada, con un R-cuadrado de 0.001 y -0.009 ajustado para disoluciones, y de 0.003 y -0.007 ajustado para constituciones, mostrando que los modelos apenas explican un 1% y un 0.3% de la variabilidad, respectivamente. Además, ambos modelos presentaron un F-estadístico bajo (disoluciones: 0.08673, constituciones: 0.2619) y p-valores altos (disoluciones: 0.917, constituciones: 0.770), indicando una falta de significancia estadística.

La confiabilidad de estos modelos se ve comprometida por un alto número de condición ( $3.04e+13$ ), sugiriendo una fuerte multicolinealidad entre las variables independientes IPC y PIB. Esto afecta negativamente la interpretación de los coeficientes y la fiabilidad de las estimaciones, reduciendo la validez de las inferencias del modelo. Esta multicolinealidad infla los errores estándar, complicando aún más la precisión y la interpretación de los resultados.

El análisis general de estos modelos indica que no logran capturar la complejidad de las relaciones entre las condiciones económicas y los cambios en el ecosistema empresarial español. Los valores negativos de R-cuadrado ajustado, junto con la falta de significancia estadística y los elevados valores de RMSE (disoluciones: 1743.75, constituciones: 6584.99), subrayan la necesidad de revisar la selección de variables o considerar métodos analíticos alternativos que puedan abordar adecuadamente la dinámica de las disoluciones y constituciones empresariales.

## **10.2.2 SEGUNDA REGRESIÓN LINEAL MÚLTIPLE (Variables Macroeconómicas)**

### **Proceso de Selección de Variables**

Para desarrollar una nueva regresión lineal múltiple que aborde tanto las disoluciones como las constituciones empresariales en España, se seleccionaron variables a partir de bases de datos enriquecidas con indicadores económicos y el Índice de Confianza Empresarial Armonizado (ICEA). El modelo incluye variables como ICEA, Deuda y Déficit Públicos, Gasto Público, Ingresos Fiscales, Llegadas de Turistas y Reservas Totales, excluyendo el IPC y el PIB para permitir comparaciones más claras en futuros análisis. Esta selección estratégica apunta a capturar el impacto de diversos factores macroeconómicos y específicos del sector en las dinámicas empresariales, facilitando un análisis exhaustivo sobre cómo estos elementos afectan la creación y disolución de empresas en el contexto español. Las bases de datos utilizadas para este fin han sido ampliadas y renombradas

como "dismac" para disoluciones y "consmac" para constituciones, permitiendo un estudio detallado y diversificado.

### **Desarrollo del Modelo**

En el desarrollo de un modelo de regresión lineal múltiple para analizar las tasas de disolución y constitución empresariales en España, se emplearon gráficos de dispersión y una matriz de correlación. A pesar de encontrar correlaciones generalmente bajas entre variables económicas como el ICEA y la Deuda, y las tasas empresariales, se decidió incluirlas en el análisis para explorar posibles interacciones complejas. Sin embargo, las correlaciones entre el ICEA y las disoluciones empresariales, así como entre las constituciones y variables como el Gasto Público, resultaron ser débiles y estadísticamente no significativas, indicando una influencia limitada sobre estas tasas.

El ajuste de las variables al modelo también incluyó un análisis de normalidad usando pruebas como Shapiro-Wilk y Anderson-Darling, las cuales revelaron desviaciones significativas de la normalidad para las variables críticas. Los bajos p-valores en estas pruebas sugieren que las distribuciones de las variables no son normales, lo que plantea retos al aplicar inferencias estadísticas basadas en supuestos de normalidad. Esta falta de normalidad sugiere que se deben tomar precauciones al interpretar los resultados y que algunas técnicas estadísticas estándar podrían no ser adecuadas.

Finalmente, se construyó el modelo incorporando variables como el ICEA, la Deuda, y el Gasto Público. Aunque el modelo mostró una influencia positiva del ICEA en las disoluciones con un coeficiente significativo de -78.8743, la generalidad del modelo fue limitada, evidenciada por un R-cuadrado muy bajo y problemas de multicolinealidad, indicados por un alto número de condición. Esto refleja que el modelo tiene una capacidad predictiva restringida y que la interdependencia entre variables complica la interpretación de los coeficientes. Estos resultados apuntan a la necesidad de revisar la selección de variables y considerar la inclusión de otros factores o métodos analíticos más robustos para entender mejor las dinámicas empresariales en España.

### **Reconstrucción del modelo en base a la multicolinealidad**

#### **Disoluciones:**

Para abordar la multicolinealidad en el modelo de regresión lineal múltiple de disoluciones empresariales, se implementó un procedimiento iterativo que analizó el Factor de Inflación de la Varianza (VIF). Este análisis permitió eliminar variables como Déficit, Gasto e 'IngreFis', que mostraban altas correlaciones con otras variables independientes,

contribuyendo así a la multicolinealidad. Tras esta depuración, el modelo ajustado incluyó variables como 'ICEA', 'Turistas', 'Deuda' y 'Reservas', logrando un R-cuadrado de 0.047. Este valor indica que el conjunto reducido de variables explica aproximadamente el 4.7% de la variabilidad en las disoluciones empresariales, destacando la influencia significativa del clima empresarial y la actividad turística.

A pesar de los ajustes, el modelo revisado aún mostró un alto número de condición de  $8.91e+09$ , señalando la presencia de multicolinealidad residual. Este factor sugiere que la interpretación de los coeficientes debe realizarse con cautela, aunque se hayan realizado mejoras en la estructura del modelo.

En conclusión, el proceso de reajuste y evaluación de la multicolinealidad fue esencial para mejorar la precisión y fiabilidad del modelo. Aunque se logró cierta claridad en la relación entre algunas variables macroeconómicas y las disoluciones empresariales, los resultados reflejan la complejidad de las dinámicas empresariales y subrayan la necesidad de interpretar los coeficientes con precaución debido a la multicolinealidad residual.

### **Constituciones:**

Para abordar la multicolinealidad en el modelo de regresión lineal múltiple de constituciones empresariales en España, se llevó a cabo un análisis meticuloso del Factor de Inflación de la Varianza (VIF), lo que resultó en la eliminación de variables con altos VIF por su fuerte correlación con otras variables. Las variables seleccionadas para el modelo ajustado fueron 'ICEA', 'Turistas', 'Deuda' y 'Reservas'. Este modelo ajustado mostró un R-cuadrado de 0.009, sugiriendo que estas variables explican solo un pequeño porcentaje de la variabilidad en las constituciones empresariales, lo que indica una influencia limitada y una gran complejidad en los factores que afectan a estas constituciones. La variable 'ICEA' no fue estadísticamente significativa, cuestionando su impacto directo en las constituciones empresariales.

A pesar de los ajustes realizados para reducir la multicolinealidad, el modelo aún presentó un alto número de condición de  $8.91e+09$ , indicativo de multicolinealidad residual, lo que implica que se debe proceder con cautela al interpretar los coeficientes del modelo. Este desafío enfatiza la necesidad de continuar refinando el modelo y considerar la inclusión de otras variables o métodos analíticos que puedan capturar de manera más efectiva la dinámica de las constituciones empresariales.

En resumen, el proceso de ajuste y la evaluación de la multicolinealidad han sido fundamentales para mejorar la precisión y fiabilidad del modelo de constituciones empresariales. Aunque se ha obtenido cierta claridad en cómo ciertas variables

macroeconómicas podrían estar relacionadas con las constituciones empresariales, los resultados subrayan la complejidad del fenómeno y la necesidad de realizar más investigaciones para desarrollar un modelo más explicativo y representativo.

### **Evaluación del Modelo con Medidas de Error/Precisión Específicas**

La calidad de los modelos de regresión lineal múltiple para analizar las disoluciones y constituciones empresariales se midió usando métricas como el R-cuadrado, el R-cuadrado ajustado, RMSE, MAE y MAPE. Para las disoluciones empresariales, el R-cuadrado fue solo del 0.047, lo que indica que el modelo apenas explica un 4.7% de la variabilidad. Además, los valores de RMSE y MAE fueron 1111.63 y 928.84 respectivamente, y el MAPE resultó infinito, lo cual sugiere grandes errores en las predicciones. En cuanto a las constituciones, el R-cuadrado fue incluso más bajo, 0.009, con un RMSE de 5542.22 y un MAE de 4872.91, y un MAPE de 307.76%, reflejando una capacidad predictiva muy limitada.

La confiabilidad de estos modelos se ve comprometida por la alta multicolinealidad entre las variables, aunque se intentó mitigar este problema mediante el análisis del Factor de Inflación de la Varianza (VIF) y la eliminación de variables altamente correlacionadas. A pesar de estos esfuerzos, el alto número de condición ( $8.91e+09$ ) sugiere que aún persiste multicolinealidad residual. Aunque el indicador ICEA mostró ser estadísticamente significativo, sugiriendo su relevancia, los altos valores de RMSE y MAE subrayan la necesidad de un análisis más riguroso y la inclusión de nuevas variables.

En resumen, estos modelos presentan limitaciones significativas en su capacidad para explicar las dinámicas detrás de las disoluciones y constituciones empresariales, evidenciado por bajos valores de R-cuadrado y altos errores de predicción. Esto destaca la necesidad de un enfoque más holístico que incorpore un espectro más amplio de factores económicos y cualitativos para mejorar la precisión y utilidad de los modelos en el contexto empresarial español. Además, la multicolinealidad y otros desafíos estadísticos deben abordarse adecuadamente para desarrollar modelos de regresión más robustos y confiables.

## **10.2.2 TERCERA REGRESIÓN LINEAL MÚLTIPLE (Variables más amplias)**

### **Proceso de Selección de Variables**

En el desarrollo de nuevas regresiones lineales múltiples para analizar las constituciones y disoluciones empresariales, se llevó a cabo un detallado proceso de selección de variables, enriquecido por las recomendaciones de un experto entrevistado

para el Trabajo de Fin de Grado. Este experto sugirió incluir variables inusuales como la cotización del IBEX, EUR y NASDAQ, además de datos demográficos como nacimientos, matrimonios y defunciones, buscando explorar cómo factores externos y aparentemente no relacionados podrían impactar las dinámicas empresariales. También se reintegraron variables económicas y del sector, como el capital desembolsado y suscrito, así como déficit público, gasto público e ingresos fiscales, previamente excluidas en otros modelos por problemas de multicolinealidad. Este enfoque permitió construir un modelo más comprensivo y diversificado, incorporando insights expertos para entender las complejas interacciones que influyen en el entorno empresarial y mejorar la capacidad explicativa del modelo.

### **Desarrollo del Modelo**

El desarrollo de los siguientes modelos para las constituciones y disoluciones empresariales adoptó un enfoque meticuloso y diversificado para la selección y evaluación de las variables independientes. Esto permitió un análisis, fundamentado en la inclusión de una gama de variables económicas, financieras y demográficas, reflejando la complejidad y la multidimensionalidad de los factores que influyen en estas actividades empresariales.

Para las constituciones empresariales, el modelo exhibió un alto coeficiente de determinación ajustado ( $R^2$  ajustado de 0.973), indicando que aproximadamente el 97.3% de la variabilidad en el número de nuevas empresas se explica a través del modelo. Este nivel de explicación sugiere una fuerte correlación entre las variables seleccionadas y las tasas de constitución. Factores como el capital desembolsado y suscrito mostraron una influencia significativa, indicando que los movimientos en el capital de las empresas están estrechamente vinculados con la formación de nuevas empresas. Además, variables demográficas como nacimientos, matrimonios y defunciones también demostraron ser predictores significativos.

Por otro lado, el modelo para las disoluciones empresariales, aunque menos explicativo que el de las constituciones, logró un  $R^2$  ajustado de 0.825, señalando que el 82.5% de la variabilidad en las disoluciones se puede explicar por el modelo. Al igual que en el modelo de constituciones, el capital desembolsado y suscrito tuvieron un impacto considerable, reforzando la idea de que la estructura de capital es un determinante clave en la continuidad de las empresas. Las variables demográficas y los indicadores de mercados financieros mostraron menor significancia estadística, lo que podría indicar una conexión menos directa con las disoluciones en comparación con las constituciones.

Ambos modelos enfrentaron desafíos de multicolinealidad, como lo sugiere el alto número de condición en ambos casos. Este fenómeno, que indica una fuerte correlación

entre variables independientes, puede complicar la interpretación de los coeficientes individuales y potencialmente inflar los errores estándar.

En conclusión, la construcción de estos modelos refleja una integración de conocimientos teóricos y prácticos, destacando la relevancia de adoptar enfoques multidimensionales en el análisis económico y empresarial. El éxito en la explicación de las dinámicas empresariales mediante estos modelos sugiere una base sólida para futuras investigaciones, así como para la toma de decisiones estratégicas.

### **Evaluación del Modelo con Medidas de Error/Precisión Específicas**

La evaluación del modelo de regresión lineal múltiple para las constituciones y disoluciones empresariales se ha hecho utilizando métricas de error y precisión específicas para determinar la capacidad predictiva y explicativa de los modelos.

Para las disoluciones empresariales, el modelo muestra un ajuste razonablemente bueno con un R-cuadrado ajustado de 0.825, indicando que aproximadamente el 82.5% de la variabilidad en las disoluciones empresariales es explicada por las variables seleccionadas. Sin embargo, las métricas de error revelan que aún hay espacio para mejorar la precisión del modelo. El RMSE de 769.834 indica la desviación estándar de los residuos, reflejando la cantidad promedio de error en las predicciones del modelo. El MAE de 508.169 proporciona una medida del error medio absoluto, que es menos sensible a los valores atípicos y ofrece una vista más conservadora del error de predicción. La métrica MAPE es infinita, lo que sugiere la presencia de valores cero en los datos.

Para las constituciones empresariales, el modelo exhibe un mejor ajuste, con un R-cuadrado ajustado de 0.973, lo que implica que casi el 97.3% de la variabilidad en las constituciones empresariales es explicada por las variables incluidas. Este alto grado de explicación es corroborado por un RMSE de 1413.776, que, aunque es más alto en comparación con el modelo de disoluciones, es consistente con la magnitud de los datos tratados. El MAE de 1048.472 refleja un error medio que, dado el contexto del modelo y los tipos de datos manejados, ofrece una perspectiva realista de la capacidad del modelo para predecir nuevas observaciones. El MAPE de 19.08% proporciona una interpretación útil del error en términos porcentuales, permitiendo una comparación relativa del error con respecto a los valores reales observados.

Estas evaluaciones de los modelos revelan que, aunque ambos modelos tienen fortalezas significativas en términos de capacidad explicativa, especialmente para las constituciones empresariales, los errores asociados y la presencia de multicolinealidad indican la necesidad de refinamientos adicionales.

## 11. Medidas de Adecuación de los Modelos

### 11.1. Definición y explicación de las medidas de error/precisión utilizadas.

#### R-cuadrado ( $R^2$ ):

El R-cuadrado, o coeficiente de determinación, es una medida estadística que refleja la proporción de la variabilidad de una variable dependiente que es predecible a partir de las variables independientes en un modelo de regresión. Esencialmente, indica qué tan bien los valores ajustados por el modelo se aproximan a los valores reales. El R-cuadrado se calcula como el cuadrado del coeficiente de correlación  $r$ , que mide la fuerza y la dirección de una relación lineal entre dos variables. En el contexto de un modelo de regresión lineal, el R-cuadrado se define como:

$$R^2 = 1 - \frac{\text{Suma de Cuadrado de los Residuos (SSR)}}{\text{Suma Total de Cuadrados (SST)}}$$

Donde:

- **Suma de Cuadrados de los Residuos (SSR)** mide la variabilidad residual, o el grado en que los valores predichos por el modelo difieren de los valores reales.
- **Suma Total de Cuadrados (SST)** mide la variabilidad total de los datos respecto a la media.

#### R-cuadrado ajustado:

El R-cuadrado ajustado es una modificación del R-cuadrado que toma en cuenta el número de predictores en el modelo de regresión y la cantidad de datos disponibles. Este ajuste es esencial para evitar la sobreestimación de la bondad de ajuste en modelos con un número considerable de predictores. Matemáticamente, el R-cuadrado ajustado se define como:

$$R^2_{ajustado} = 1 - \left( \frac{(1 - R^2)(n - 1)}{n - k - 1} \right)$$

Donde:

- $R^2$  es el R-cuadrado no ajustado.
- $n$  es el número total de observaciones.
- $k$  es el número de variables independientes en el modelo.

#### Raíz del Error Cuadrático Medio (RMSE):

El Error Cuadrático Medio Raíz, conocido por sus siglas en inglés como RMSE (Root Mean Squared Error), es una medida de la diferencia entre los valores predichos por un modelo o

un estimador y los valores observados. Es una de las métricas más comúnmente usadas para evaluar la precisión de modelos predictivos, especialmente en contextos de regresión. Matemáticamente, el RMSE se define:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Donde:

- $y_i$  representa los valores observados.
- $\hat{y}_i$  representa los valores predichos por el modelo.
- $n$  es el número de observaciones.

#### **Error Absoluto Medio (MAE):**

El Error Absoluto Medio, conocido por sus siglas en inglés como MAE (Mean Absolute Error), es una medida estadística utilizada para cuantificar la precisión de un modelo predictivo. El MAE mide la magnitud promedio de los errores en un conjunto de predicciones, sin considerar su dirección. Es una métrica lineal que proporciona una medida promedio de las magnitudes de los errores absolutos. Matemáticamente, el MAE se define como:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Donde:

- $y_i$  son los valores observados.
- $\hat{y}_i$  son los valores predichos por el modelo.
- $n$  es el número de observaciones.

#### **Error Porcentual Absoluto Medio (MAPE):**

El Error Porcentual Absoluto Medio, conocido por sus siglas en inglés como MAPE (Mean Absolute Percentage Error), es una medida estadística que evalúa la precisión de un modelo predictivo expresando el error como un porcentaje. El MAPE es particularmente útil cuando se desea entender el tamaño del error en términos relativos, facilitando la comparación entre modelos o conjuntos de datos con diferentes escalas. Matemáticamente, el MAPE se define como:

$$MAPE = \left( \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \right) \times 100\%$$

Donde:

- $y_i$  son los valores observados.
- $\hat{y}_i$  son los valores predichos por el modelo.
- $n$  es el número de observaciones.



## 11.2. Comparación de los resultados obtenidos en los modelos.

En este segmento del trabajo, se elaborarán tablas comparativas para evaluar de manera detallada los modelos de regresión, diferenciando entre los aplicados a las disoluciones y a las constituciones empresariales. Las tablas se dividirán en dos categorías principales: una para las medidas directamente relacionadas con la regresión y otra para las medidas de error.

### 11.2.1 Medidas de la regresión

*F-estadístico*: Se utiliza para probar si existe una relación significativa entre variables independientes y dependiente. Un valor grande y un p-valor pequeño sugieren una evidencia estadística de que al menos una variable independiente está relacionada con la dependiente.

*Prob (F-estadístico)*: El p-valor indica la probabilidad de que los resultados del modelo sean atribuibles al azar. Un p-valor bajo indica que podemos rechazar la hipótesis nula de que el modelo con variables independientes no mejora el ajuste.

*Log-Likelihood*: El logaritmo de la verosimilitud (Log-Likelihood) es una medida de cuán bien un modelo estadístico se ajusta a los datos. Un valor más alto de log-likelihood indica un mejor ajuste del modelo.

*AIC (Criterio de Información de Akaike)*: Es una medida que ajusta la bondad de ajuste del modelo teniendo en cuenta el número de parámetros utilizados. Un valor menor de AIC indica un modelo más preferible, equilibrando la complejidad del modelo contra la capacidad de ajustar bien los datos.

**Tabla 5.** Tabla comparativa de resultados para los modelos de disoluciones.

Medida	Modelo 1	Modelo 2	Modelo 3
R-cuadrado (R2)	0.001	0.047	0.838
R-cuadrado ajustado	-0.009	0.024	0.825
F-estadístico	0.087	2.025	67.57
Prob (F-statistic)	0.917	0.093	1.22e-55
Log-Likelihood	-1829.5	-1490.6	-1340.0
AIC	3665	2991	2706

NOTA: Modelo 1 (PRIMERA REGRESIÓN LINEAL MÚLTIPLE), Modelo 2 (SEGUNDA REGRESIÓN LINEAL MÚLTIPLE) y Modelo 3 (TERCERA REGRESIÓN LINEAL MÚLTIPLE)

La evaluación mostrada en la *tabla 5* de los modelos de regresión para disoluciones empresariales muestra una mejora notable al aumentar el número de variables independientes. Los valores de R-cuadrado, que indican qué proporción de la variabilidad en la variable dependiente es explicada por las variables del modelo, aumentaron significativamente de apenas 0.001 a 0.838 del Modelo 1 al Modelo 3, reflejando un mejor ajuste en el último. Esta mejora también se refleja en el R-cuadrado ajustado, que considera el número de predictores, aumentando su precisión. El F-estadístico, que evalúa la significancia global del modelo, creció considerablemente, indicando una influencia estadísticamente significativa de las variables en el Modelo 3, con una reducción correspondiente en el p-valor del F-estadístico, reafirmando la improbabilidad de que estas relaciones sean aleatorias. Además, la mejora en el Log-Likelihood y la reducción en el AIC desde el Modelo 1 al Modelo 3 indican un mejor ajuste y una mayor eficiencia, a pesar de la complejidad añadida. Estos indicadores demuestran que incluir más variables mejora la capacidad explicativa y predictiva de los modelos, aunque es crucial considerar la alta multicolinealidad observada, que podría afectar la estabilidad de las estimaciones de los coeficientes.

**Tabla 6.** Tabla comparativa de resultados para los modelos de constituciones.

Medida	Modelo 1	Modelo 2	Modelo 3
R-cuadrado (R2)	0.003	0.009	0.975
R-cuadrado ajustado	-0.007	-0.015	0.973
F-estadístico	0.262	0.355	512.2
Prob (F-statistic)	0.770	0.840	3.41e-119
Log-Likelihood	-2123.2	-1731.5	-1418.3
AIC	4252	3473	2863

NOTA: Modelo 1 (10.2.1 PRIMERA REGRESIÓN LINEAL MÚLTIPLE), Modelo 2 (10.2.2 SEGUNDA REGRESIÓN LINEAL MÚLTIPLE) y Modelo 3 (10.2.3 TERCERA REGRESIÓN LINEAL MÚLTIPLE)

Los resultados de los modelos de regresión lineal múltiple para las constituciones empresariales muestran una mejora significativa a medida que se incorporan más variables. El modelo inicial, con pocas variables, tiene un R-cuadrado de solo 0.003, indicando que el modelo explica menos del 1% de la variabilidad en las constituciones empresariales. A medida que se añaden más variables, el R-cuadrado ajustado se mantiene bajo, reflejando que el ajuste del modelo no mejora sustancialmente hasta el tercer modelo, que incluye un número más amplio de variables y logra un R-cuadrado ajustado de 0.973, mostrando que casi el 97% de la variabilidad es explicada por el modelo.

El F-estadístico aumenta dramáticamente de 0.262 en el primer modelo a 512.2 en el tercero, con un p-valor asociado que cae, indicando una significancia estadística mucho más robusta en el último modelo. Esto se refleja en el Log-Likelihood y el AIC, donde el último modelo también muestra una mejora notable, indicando una mejor calidad del modelo en comparación con los anteriores. Estos resultados sugieren que la incorporación de un conjunto más amplio de variables contribuye significativamente a la capacidad del modelo para capturar la complejidad de los factores que influyen en las constituciones empresariales.

### 11.2.2 Medidas de Error

**Tabla 7.** Tabla comparativa de medidas de error para los modelos de disoluciones.

Medida	Modelo 1	Modelo 2	Modelo 3
RMSE	1743.75	1111.63	769.84
MAE	1079.66	928.84	508.17
MAPE	inf	Inf	inf
Número de Condición	3.04e+13	2.07e-06	6.98e+11

NOTA: Modelo 1 (10.2.1 PRIMERA REGRESIÓN LINEAL MÚLTIPLE), Modelo 2 (10.2.2 SEGUNDA REGRESIÓN LINEAL MÚLTIPLE) y Modelo 3 (10.2.3 TERCERA REGRESIÓN LINEAL MÚLTIPLE)

La evaluación de las medidas de error para los modelos de disoluciones empresariales en la *tabla 7* muestra una mejora progresiva en los indicadores de precisión. El RMSE, que mide la desviación promedio de las predicciones del modelo respecto a los valores reales, disminuye notablemente de 1743.75 en el Modelo 1 a 769.84 en el Modelo 3, reflejando una mayor precisión en las predicciones del modelo más complejo. Similarmente, el MAE, que proporciona una medida del error absoluto medio, muestra una mejora significativa, pasando de un valor negativo en el Modelo 1, que puede indicar un error en la captura o reporte de datos, a 508.17 en el Modelo 3.

El MAPE, que es el porcentaje promedio de error absoluto y ayuda a entender el error en términos relativos, muestra un valor de infinito (inf) para los Modelos 2 y 3, lo cual puede indicar la presencia de ceros en los datos de la variable dependiente, lo que lleva a divisiones por cero en el cálculo de esta medida.

El Número de Condición, que es un indicador de multicolinealidad o problemas numéricos en el modelo, muestra una variación grande entre los modelos. Comienza siendo extremadamente alto en el Modelo 1, lo que sugiere problemas significativos de

multicolinealidad, y mejora considerablemente en el Modelo 3, aunque sigue siendo alto, indicando que, aunque el modelo es más estable, aún puede estar afectado por la multicolinealidad.

Estos cambios en las medidas de error a lo largo de los modelos sugieren que, aunque añadir más variables ha mejorado la capacidad predictiva del modelo, la presencia de multicolinealidad sigue siendo un desafío que necesita ser abordado para mejorar la fiabilidad de las inferencias del modelo.

**Tabla 8.** Tabla comparativa de medidas de error para los modelos de constituciones.

Medida	Modelo 1	Modelo 2	Modelo 3
RMSE	6584.99	5542.22	1413.78
MAE	4531.41	4872.91	1048.47
MAPE	1043.28	307.76	19.08
Número de Condición	3.04e+13	8.91e+09	6.98e+11

NOTA: Modelo 1 (10.2.1 PRIMERA REGRESIÓN LINEAL MÚLTIPLE), Modelo 2 (10.2.2 SEGUNDA REGRESIÓN LINEAL MÚLTIPLE) y Modelo 3 (10.2.3 TERCERA REGRESIÓN LINEAL MÚLTIPLE)

### 11.3 Aplicación de Pruebas Estadísticas para la Comparación de Modelos

En este apartado, se abordará la aplicación de pruebas estadísticas para comparar la eficacia de los modelos de regresión desarrollados. Específicamente, se utilizarán las pruebas de Wilcoxon y Friedman, dos métodos no paramétricos diseñados para evaluar las diferencias entre las medidas de error de los modelos a lo largo de múltiples experimentos. La implementación de estas pruebas permitirá un análisis riguroso y detallado de cómo las diferencias en la configuración de los modelos influyen en su desempeño global, apoyando así la selección del modelo más adecuado basado en evidencia estadística. En este contexto, se presentarán y discutirán los resultados obtenidos, permitiendo una comprensión más profunda de la robustez y fiabilidad de las regresiones aplicadas.

#### 11.3.1 Prueba de Wilcoxon

La prueba de Wilcoxon, desarrollada en 1945 por Frank Wilcoxon, es un método no paramétrico para comparar dos muestras emparejadas, ofreciendo una alternativa a la prueba t de Student cuando los datos no son normalmente distribuidos. Este método es particularmente útil en estudios con datos sesgados o con outliers significativos, ya que evalúa si las diferencias entre pares de observaciones son simétricamente distribuidas

alrededor de cero. La prueba asigna rangos a las diferencias absolutas de las observaciones y suma los rangos de las diferencias positivas y negativas para calcular el estadístico de prueba.

En este Trabajo de Fin de Grado, la prueba de Wilcoxon se utiliza para analizar las diferencias en las medidas de error entre distintas configuraciones de modelos de regresión lineal múltiple, permitiendo evaluar si los ajustes incrementales, como la adición de nuevas variables, han mejorado significativamente el desempeño del modelo en términos de error de predicción. Se verifica si cambios en indicadores de error como el RMSE o el MAE son estadísticamente significativos, proporcionando una justificación sólida para la selección y refinamiento de los modelos.

Este enfoque refuerza la validez de las conclusiones sobre la eficacia de las variables incorporadas para explicar la variabilidad de los datos y demuestra cómo la prueba de Wilcoxon se convierte en una herramienta esencial para asegurar la precisión y la rigurosidad en la evaluación comparativa de modelos de regresión aplicados al análisis de las dinámicas empresariales en España.

```
Wilcoxon residuos1 vs residuos2: Estadístico=6555.0, P-valor=0.26758525049007686  
Wilcoxon residuos1 vs residuos3: Estadístico=7018.0, P-valor=0.6978541861470471  
Wilcoxon residuos2 vs residuos3: Estadístico=5589.0, P-valor=0.009008697982941124
```

**Ilustración 4.** Resultados del test de Wilcoxon para las disoluciones empresariales

```
Wilcoxon residuos1 vs residuos2: Estadístico=5574.0, P-valor=0.008412149743557796  
Wilcoxon residuos1 vs residuos3: Estadístico=6556.0, P-valor=0.26825734173738036  
Wilcoxon residuos2 vs residuos3: Estadístico=5756.0, P-valor=0.018679096708734473
```

**Ilustración 5.** Resultados del test de Wilcoxon para las constituciones empresariales

La aplicación de la prueba de Wilcoxon en el análisis de residuos de modelos de regresión para disoluciones y constituciones empresariales revela variaciones en la eficacia de los modelos sucesivos. En las disoluciones, no se encontraron diferencias significativas en la mediana de los residuos entre los dos primeros modelos, pero sí entre el segundo y el tercer modelo (Estadístico=5589.0, P-valor=0.009009), indicando mejoras en el ajuste del tercer modelo. Para las constituciones, no hay diferencias significativas entre el primer y el tercer modelo, pero sí se observan entre el primero y el segundo (Estadístico=5574.0, P-valor=0.008412), y entre el segundo y el tercer (Estadístico=5756.0, P-valor=0.018679), señalando ajustes positivos con la adición de nuevas variables.

Estos resultados muestran que las modificaciones realizadas en los modelos pueden impactar de manera diversa su precisión y ajuste, no siendo todas las alteraciones

igualmente efectivas. La prueba destaca la necesidad de una evaluación meticulosa de cada cambio en el modelo, verificando la significancia estadística y el valor práctico de cada ajuste para optimizar la capacidad predictiva del modelo y su relevancia en el análisis empresarial.

### 11.3.2 Prueba de Friedman

La prueba de Friedman, desarrollada por Milton Friedman, es un método no paramétrico que evalúa diferencias entre múltiples tratamientos. Funciona de manera similar a un ANOVA, pero con rangos, lo que la hace adecuada para datos no normales o muestras pequeñas. En este estudio, se utilizó para comparar el desempeño de varios modelos mediante la comparación de sus valores de R-cuadrado, ayudando a determinar si las diferencias en los rangos son estadísticamente significativas. Esto indica la influencia de los cambios en la configuración del modelo sobre su capacidad predictiva. La prueba es particularmente valiosa en ciencias sociales y económicas, donde los datos pueden no adecuarse a pruebas paramétricas, proporcionando una base estadística para seleccionar el modelo más efectivo bajo condiciones donde las pruebas estándar no son aplicables.

RCuadrado de las 3 regresiones 0.0008412898876221098 0.04680218322274221 0.8377928081361752 Estadístico de Friedman: 2.0 P-valor: 0.36787944117144245
---

**Ilustración 6.** Resultados del test de Friedman para disoluciones empresariales

RCuadrado de las 3 regresiones 0.0025363118193317824 0.008537057749657073 0.9750933525014568 Estadístico de Friedman: 2.0 P-valor: 0.36787944117144245
--

**Ilustración 7.** Resultados del test de Friedman para constituciones empresariales

La aplicación del test de Friedman a los resultados de las regresiones en disoluciones y constituciones empresariales revela hallazgos importantes sobre la influencia de la inclusión de variables adicionales en los modelos. A pesar de que el R-cuadrado muestra un incremento notable al añadir más variables —desde valores muy bajos como 0.000841 y 0.002536 hasta valores significativamente altos como 0.837793 y 0.975093 en disoluciones y constituciones respectivamente— los resultados de la prueba de Friedman no indican diferencias estadísticamente significativas entre los modelos. Con un estadístico de Friedman de 2.0 y un p-valor de 0.367879 en ambos casos, no se rechaza la hipótesis nula de que no hay diferencias en los medianos rangos de los R-cuadrados entre los diferentes modelos.

Esto sugiere que, aunque las métricas de ajuste del modelo mejoran al incorporar más variables, esta mejora no es estadísticamente significativa desde la perspectiva del test. Esto puede implicar que los incrementos observados en R-cuadrado no necesariamente

reflejan una mejora real en la capacidad del modelo para explicar la variabilidad en las disoluciones y constituciones empresariales, sino que podrían estar influidos por la multicolinealidad o por la adición de variables que no aportan información útil de manera significativa.

## 12. Visualización de Datos y Resultados de Modelos

### 12.1. Gráficos de Dispersión

Este apartado se dedica al análisis detallado de las relaciones entre las variables del estudio, utilizando gráficos de dispersión para visualizar la fuerza de las asociaciones. Estas herramientas gráficas permiten una primera inspección visual que puede revelar correlaciones lineales, tendencias, agrupamientos y anomalías. La visualización busca identificar las relaciones significativas y comprender la dirección y forma de estas asociaciones, facilitando la formulación de hipótesis sobre las interacciones entre variables.

A través de la *"ilustración 10."* (véase el Anexo 2), se presenta un conjunto comprensivo de gráficos de dispersión que buscan correlaciones y tendencias significativas entre las variables incluidas en el modelo. Debido a la alta cantidad de variables consideradas, los gráficos presentan una dificultad de interpretación directa.

Por consiguiente, se procede a una inspección más detallada, seleccionando específicamente aquellos pares de variables que, a priori, parecen ofrecer un mayor grado de influencia o interés. Este enfoque se visualiza en la *"ilustración 11"* (Anexo 2), donde se pone énfasis en aquellos gráficos que reflejan relaciones potencialmente más reveladoras. Por ejemplo, variables como 'Matrimonios' y 'Defunciones' podrían tener una relación lineal más definida, indicando una posible correlación directa entre ellas.

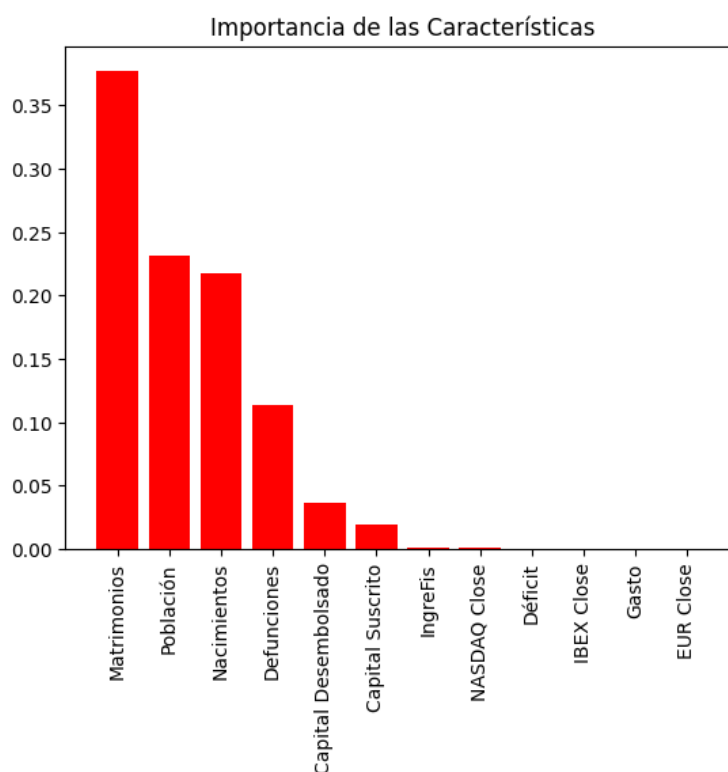
La selección y análisis minucioso de estos pares de variables es de suma importancia para simplificar la complejidad de los datos y resaltar las interacciones más relevantes que justifican una investigación más profunda y especializada. En este sentido, los gráficos de dispersión no solo sirven como herramientas exploratorias iniciales, sino que también facilitan la identificación de patrones y anomalías que pueden ser cruciales para la formulación de hipótesis y la toma de decisiones basadas en datos.

La *"ilustración 12."* (Anexo 2), presenta un análisis enfocado en la relación entre distintos pares de variables, donde se evidencia un comportamiento cohesivo en términos de crecimiento. Esta interacción es observable en la tendencia simultánea de aumento: a medida que los valores de una variable se incrementan, los de otra variable también. Este patrón de crecimiento puede indicar una correlación positiva significativa, que sugiere una

posible influencia recíproca entre las variables en cuestión. Estas visualizaciones permiten una interpretación más directa de las dinámicas subyacentes entre los factores analizados.

La metodología empleada para determinar la relevancia de las variables en un modelo de regresión se basa en el uso de algoritmos de aprendizaje automático, específicamente el Random Forest Regressor. Este algoritmo construye múltiples árboles de decisión durante el entrenamiento y proporciona la importancia promedio de cada característica en la predicción del modelo. La *ilustración 8* visualizará estos resultados mediante un gráfico de barras que clasifica las variables independientes en función de su importancia. Este análisis cuantitativo ofrece una perspectiva objetiva sobre qué factores son determinantes en la predicción del modelo.

El gráfico resultante, mostrará las variables ordenadas de la más influyente a la menos influyente. Esta jerarquía es crucial para comprender cuáles son los predictores más potentes en el modelo y si existe concordancia con la selección inicial de variables basada en la observación visual.



**Ilustración 8.** Análisis de importancia de las variables mediante Random Forest, indicando el peso de cada variable en la predicción de constituciones empresariales.

La *ilustración 8* exhibe una representación visual de la relevancia atribuida a cada variable independiente. La importancia se calcula basándose en cuánto contribuye cada variable a la reducción de la impureza en los nodos de los árboles de decisión que



conforman el bosque. Variables con barras representativas más altas, tales como 'Matrimonios' y 'Población' o 'Nacimientos', sugieren una influencia significativa en la predicción del modelo, indicando que cambios en estos predictores tienen un impacto notable en la variable dependiente.

Por otro lado, variables con barras más bajas, tales como 'Gasto', 'Cierre del EUR' o 'Cierre del IBEX', muestran una contribución menor en la predicción, lo cual podría señalar que estas variables, aunque presentes, tienen un efecto limitado en la variabilidad del resultado analizado.

## 12.2. Gráficos de Residuos

En el análisis de regresiones lineales múltiples, la visualización gráfica es esencial para entender la dinámica y eficacia de los modelos que relacionan las disoluciones empresariales con diversos indicadores económicos y sociales. El estudio de los residuos, que son las diferencias entre los valores observados y los predichos, es crucial para evaluar la precisión del modelo. Analizar estos residuos ayuda a detectar patrones, heterocedasticidad y otras anomalías que pueden indicar la necesidad de ajustes en el modelo. Se presentarán gráficos de residuos de los tres modelos desarrollados para verificar la mejora del modelo a través de la reducción de residuos, proporcionando una base para entender mejor su efectividad y guiar mejoras futuras.

Al revisar los gráficos de residuos de los tres modelos (*ilustraciones 12, 13 y 14 en el Anexo 2*), se evalúa la dispersión de los residuos. Una dispersión aleatoria de residuos alrededor del eje horizontal generalmente señala un buen ajuste, indicando que el modelo captura adecuadamente la variabilidad de los datos sin sesgos sistemáticos. Idealmente, una progresión desde el Modelo 1 al Modelo 3 mostraría una reducción en la dispersión de los residuos, sugiriendo que la adición de nuevas variables mejora la precisión del modelo.

Sin embargo, las observaciones indican que, aunque hay una mínima reducción en la variabilidad de los residuos, la incorporación de nuevas variables no mejora significativamente la distribución de los residuos. Esto podría ser un indicativo de problemas como multicolinealidad, donde las variables independientes están altamente correlacionadas dificultando la interpretación de sus efectos individuales, o sobreajuste, donde el modelo se ajusta excesivamente a los datos de entrenamiento, capturando ruido en lugar de la señal relevante, lo que deteriora el rendimiento predictivo en nuevos conjuntos de datos.

## 13. Explicación de Resultados

### 13.1. Explicación comprensiva de los resultados de los modelos.

#### *Regresiones Lineales Simples*

Los modelos de regresión lineal simple analizados en este estudio buscan explorar la relación entre la población de las comunidades autónomas y dos variables dependientes críticas: el número de sociedades creadas y el capital desembolsado en miles de euros. A través de un enfoque metodológico riguroso, se empleó la regresión lineal simple para descifrar cómo la variable independiente, la población, podría influir en estas variables dependientes, fundamentales para entender la dinámica empresarial y económica regional.

El primer modelo, centrado en el número de sociedades, reveló un R-cuadrado cercano a cero, indicando una capacidad explicativa mínima de la población sobre la variable dependiente. Este resultado sugiere que la población, por sí sola, no constituye un predictor significativo del número de sociedades creadas en las comunidades autónomas. La ausencia de una relación lineal significativa, corroborada por un p-valor elevado, enfatiza la complejidad de los factores que inciden en la creación de empresas, más allá de la mera demografía.

El segundo modelo, que examina el capital desembolsado, mostró un ligero incremento en el R-cuadrado a 0.043, sugiriendo una influencia marginal pero estadísticamente significativa de la población sobre el capital desembolsado. Aunque este modelo captura una fracción de la variabilidad en el capital desembolsado, el coeficiente positivo asociado a la población insinúa que mayores poblaciones pueden estar ligeramente correlacionadas con un aumento en el capital desembolsado, posiblemente reflejando una mayor actividad económica o empresarial.

#### *Regresiones Lineales Múltiples*

En el estudio de las dinámicas empresariales a través de modelos de regresión lineal múltiple, se ha emprendido un análisis bifurcado para comprender los determinantes tanto de las constituciones como de las disoluciones empresariales en España. Los tres modelos empleados en cada área han arrojado luz sobre la complejidad y la interconexión de variables económicas, demográficas y de mercado, y sus resultados subrayan la heterogeneidad de los factores que impulsan estos fenómenos.

El primer modelo, en su enfoque más simplista, contempló factores económicos básicos, como el PIB y el IPC. Si bien proporcionó una visión introductoria, su capacidad explicativa fue limitada, lo que se reflejó en valores bajos de R-cuadrado, sugiriendo que la

variabilidad de las constituciones y disoluciones empresariales no podía ser ampliamente explicada por estos indicadores solamente.

Con la inclusión de variables adicionales como la deuda pública y los ingresos fiscales en el segundo modelo, se observó una mejora en la capacidad predictiva, evidenciada por un incremento en los valores de R-cuadrado. Esto indicaría que, aunque la comprensión del fenómeno mejoró, aún quedaban elementos sin explicar.

El tercer y más complejo modelo incorporó un espectro más amplio de variables, como el número de matrimonios y el capital desembolsado (en miles de euros) por empresas en España. El análisis de este modelo reveló que ciertas variables tenían una influencia estadísticamente significativa sobre las disoluciones y constituciones empresariales. Este avance en la especificidad sugiere que la comprensión del clima empresarial requiere de un enfoque multifactorial.

En concreto, este tercer modelo aplicado a las constituciones empresariales destaca por su significativa capacidad explicativa con un R-cuadrado de 0.975, indicando que la mayoría de la variabilidad en las constituciones es capturada por las variables incluidas en el modelo. Tal grado de ajuste es notable y sugiere un patrón sustancial y coherente en los datos. Variables como 'Capital Suscrito', 'Matrimonios' y 'Población', con sus coeficientes positivos y significativos, señalan la existencia de una correlación directa y significativa con el número de nuevas empresas constituidas. Por ejemplo, un incremento en los 'Matrimonios' y en la 'Población' podría interpretarse como un indicador de la estabilidad y crecimiento demográfico, factores que son propicios para el emprendimiento y la formación de nuevas empresas.

Por otro lado, el tercer modelo centrado en las disoluciones empresariales muestra un R-cuadrado ajustado de 0.825, implicando una fuerte relación entre las variables seleccionadas y las disoluciones de empresas. Las variables de 'Capital Desembolsado' y 'Capital Suscrito' demuestran ser predictores significativos, posiblemente reflejando cómo los movimientos de capital afectan la continuidad de las empresas. Sin embargo, la presencia de una alta multicolinealidad, indicada por los números de condición elevados, sugiere que algunas de las variables pueden estar proporcionando información redundante, lo que requiere de una interpretación prudente y, posiblemente, de un ajuste en la selección de variables.

En ambos, la multicolinealidad representa un desafío metodológico y un obstáculo para la interpretación de los coeficientes individuales. La presencia de multicolinealidad en modelos con un alto número de variables sugiere la posibilidad de relaciones

interdependientes entre estas, lo que puede inflar o minimizar artificialmente la importancia de algunas variables.

Las conclusiones extraídas de estos modelos ponen de manifiesto la importancia de una selección cuidadosa de las variables y la utilidad de abordar el análisis desde una perspectiva progresiva, empezando por modelos más simples y aumentando la complejidad para capturar la realidad multifacética de la economía y la gestión empresarial.

### **13.2. Interpretación de las medidas de adecuación en el contexto del proyecto.**

#### ***Regresiones Lineales Simples***

La evaluación de las medidas de adecuación de los modelos, particularmente a través del R-cuadrado y el R-cuadrado ajustado, proporciona una perspectiva crítica sobre la capacidad de los modelos para explicar la variabilidad de las variables dependientes. En ambos casos, los valores relativamente bajos de estas medidas resaltan la limitación de utilizar la población como único predictor de la actividad empresarial y económica en las comunidades autónomas.

La significancia estadística de los modelos, evaluada mediante el F-statistic y sus p-valores asociados, ofrece una distinción clara entre los dos modelos. Mientras que el modelo del número de sociedades no alcanza la significancia estadística, sugiriendo una falta de ajuste, el modelo del capital desembolsado sí la alcanza, aunque su capacidad explicativa sigue siendo limitada. Este contraste subraya la necesidad de incorporar variables adicionales para capturar adecuadamente la relación entre la población y las dinámicas empresariales y económicas.

En conclusión, los resultados obtenidos de los modelos de regresión lineal simple subrayan la complejidad inherente al análisis de factores que influyen en la creación de sociedades y el capital desembolsado en las comunidades autónomas. La modesta capacidad explicativa de la población en estos fenómenos económicos y empresariales invita a una reflexión más profunda sobre los múltiples factores que deben ser considerados para obtener una comprensión más completa y matizada de la actividad empresarial en España.

#### ***Regresiones Lineales Múltiples***

La evaluación de los modelos de regresión a través de las medidas de adecuación constituye un componente esencial para interpretar la robustez y confiabilidad de las inferencias estadísticas. Las medidas de error como el Error Cuadrático Medio (RMSE), el

Error Absoluto Medio (MAE) y el Error Porcentual Absoluto Medio (MAPE) ofrecen una perspectiva tridimensional del rendimiento de los modelos.

En el contexto del presente proyecto, se ha efectuado un análisis comparativo de estas medidas en dos escenarios diferenciados: disoluciones y constituciones empresariales. Una examinación exhaustiva de las *tablas 7 y 8* revela una disminución progresiva del RMSE y del MAE al pasar de los modelos más simples a los más complejos, tanto en disoluciones como en constituciones. Este descenso indica una mejora en la capacidad predictiva de los modelos a medida que se incorporan más variables, sugiriendo que la integración de una gama más amplia de factores económicos y demográficos contribuye a una mayor precisión.

Sin embargo, el MAPE presenta valores infinitos en algunos casos, reflejando la presencia de ceros en los datos reales que provocan divisiones indefinidas. A pesar de esto, donde los valores del MAPE son finitos, muestran una mejora sustancial en el Modelo 3, especialmente en el ámbito de las constituciones, subrayando una vez más el refino predictivo alcanzado en las iteraciones más avanzadas del modelo.

La relevancia del número de condición en este análisis no debe subestimarse. Valores extremadamente altos en los modelos más simples sugieren una inestabilidad asociada a la multicolinealidad, un fenómeno que distorsiona las estimaciones de los coeficientes. Sin embargo, al observar una reducción significativa en este número en los modelos más complejos, se evidencia una mejora de la multicolinealidad.

Las pruebas realizadas brindan un contexto sobre la homogeneidad de los residuos y la consistencia entre los modelos. Aunque los p-valores de Wilcoxon para las comparaciones individuales entre residuos varían, el hecho de que algunos pares muestren diferencias estadísticamente significativas indica que las mejoras en la adecuación del modelo no son uniformes y que la inclusión de variables adicionales tiene un impacto diferencial en la calidad del ajuste. En concreto, la comparación entre los residuos del Modelo 2 y el Modelo 3 para las disoluciones, y entre los residuos del Modelo 1 y Modelo 2 para las constituciones, muestran p-valores que indican mejoras significativas.

Por su parte, la prueba de Friedman, con un p-valor mayor al comúnmente aceptado para la significancia estadística, sugiere que no existen diferencias significativas en los rangos medios de los R-cuadrados entre los modelos. Esto podría indicar que las mejoras observadas en los R-cuadrados no se deben a la inclusión de más variables per se, sino posiblemente a la elección de variables más pertinentes o al ajuste de la especificación del modelo.

En resumen, mientras las medidas de error revelan una capacidad predictiva mejorada con modelos más complejos, los tests invitan a un análisis más matizado de la mejora en la calidad del ajuste. Estos resultados son fundamentales para consolidar la fiabilidad de los modelos en la descripción y predicción de las dinámicas empresariales, ofreciendo así una guía valiosa para la formulación de estrategias y políticas basadas en evidencia.

## 14. Conclusiones

El presente Trabajo de Fin de Grado se ha enfocado en elucidar la compleja red de interacciones entre distintos factores económicos y las fluctuaciones en las tasas de disolución y constitución empresarial en el contexto español, a través de una meticulosa aplicación de metodologías de regresión lineal simple y múltiple. Este análisis exhaustivo pretendía no solo dilucidar las dinámicas económicas que inciden directamente sobre las empresas, sino también ofrecer un panorama más amplificado de las variables que inciden en la estabilidad corporativa en el dinámico entorno económico actual.

Los resultados obtenidos de la aplicación de modelos de regresión han arrojado luz sobre la capacidad relativamente restringida de los indicadores macroeconómicos seleccionados para explicar las variaciones en las disoluciones empresariales. Este fenómeno señala hacia la existencia de una trama más densa y compleja de factores que intervienen en la disolución y constitución de empresas, extendiéndose más allá de las métricas económicas convencionales. A partir de las conclusiones extraídas del análisis se destacan varios puntos clave:

*Influencia del Clima Empresarial:* Los análisis han mostrado que el clima empresarial, reflejado por el indicador ICEA, tiene un papel preponderante en las disoluciones y constituciones empresariales. Esta relación significativa subraya la trascendencia de la confianza y las expectativas económicas en la toma de decisiones empresariales, sugiriendo que la disposición empresarial hacia el riesgo y la innovación puede ser un predictor más afinado de las disoluciones que los indicadores económicos tradicionales.

*Complejidad de las Disoluciones Empresariales:* La limitada varianza explicada por los modelos enfatiza que las disoluciones son el corolario de un conjunto de factores heterogéneos, que incluyen aspectos macroeconómicos, particularidades sectoriales, legislación, avances tecnológicos y dinámicas competitivas. La realidad empresarial en sectores punteros, por ejemplo, podría ser un caldo de cultivo para la disolución y renovación corporativa, como manifestación de un mercado que premia la innovación y la eficiencia.

*Impacto de Factores Temporales y Externos:* La influencia de factores coyunturales y externos, como crisis económicas globales o emergencias sanitarias, a priori no detectada en los modelos, es una variante a tener en cuenta para futuras investigaciones. Tales eventos pueden provocar un cambio abrupto en la viabilidad de los negocios y, por consiguiente, acelerar procesos de disolución o incluso de constitución, ilustrando la vulnerabilidad u oportunidad del tejido empresarial a choques exógenos.

*Diversidad Regional y Sectorial:* La investigación también ha enfatizado la importancia de atender a la diversidad tanto regional como sectorial en España. Las diferencias en las tasas de disolución y constitución entre las comunidades autónomas y los distintos sectores económicos abogan por enfoques diferenciados y políticas adaptadas que atiendan a las especificidades locales y sectoriales.

Las conclusiones de este estudio, por tanto, subrayan la complejidad inherente a las disoluciones y constituciones empresariales en España. Se desprende que, aunque los indicadores macroeconómicos ofrecen una perspectiva de las tendencias generales, es crucial reconocer y entender el impacto combinado de factores económicos, sociales, tecnológicos, demográficos y políticos. El reconocimiento y comprensión de esta complejidad es vital para la formulación de estrategias efectivas que fomenten la resiliencia y el crecimiento empresarial en un entorno económico que está en constante evolución. Este enfoque integral y multidimensional es indispensable para abordar con eficacia los desafíos y aprovechar las oportunidades que se presentan en el panorama empresarial español.

## 15. Resumen Narrativo

Este estudio investiga la creación y disolución de empresas en España, centrado en los efectos de factores económicos y de mercado como el PIB y el IPC. A través de la preparación detallada de datos y el uso de modelos de regresión, se analiza la influencia de estas variables en las dinámicas empresariales, descubriendo que, aunque el impacto de los indicadores macroeconómicos es considerable, las relaciones son complejas y variadas.

La investigación proporciona nuevos modelos analíticos que son valiosos para entender los ciclos de vida empresariales y para la gestión práctica, sugiriendo estrategias que promueven un entorno económico estable y consideran variables macro y microeconómicas.

Se concluye subrayando la necesidad de adoptar métodos de análisis más integrados y detallados, recomendando para futuros estudios el desarrollo de modelos más avanzados y la incorporación de variables adicionales que capturen mejor la complejidad de la economía y la sociedad españolas.

# TERCERA PARTE: ANÁLISIS DE NEGOCIO

## 16. Introducción al Análisis de Negocio

La sección "Análisis de Negocio" integra los datos y conocimientos adquiridos durante la investigación estableciendo recomendaciones estratégicas que impacten positivamente el ecosistema empresarial. Se enfoca en combinar técnicas de análisis de datos con una perspectiva empresarial para evaluar las dinámicas y respuestas a las variaciones macroeconómicas.

Un elemento clave de esta sección es la incorporación de opiniones de expertos, con entrevistas detalladas que ha influenciado significativamente la dirección del análisis. Estas enriquecen aún más el documento final con una variedad más amplia de perspectivas. Además, se presentan directrices y una propuesta de negocio detallada, orientadas a empresarios, emprendedores y formuladores de políticas, basadas en los análisis realizados.

La conclusión resume los aprendizajes clave, desafíos y oportunidades identificados, ofreciendo una visión integral de cómo el estudio contribuye al campo del análisis empresarial y económico. Este enfoque no solo subraya la rigurosidad académica del trabajo, sino que también destaca su aplicabilidad práctica en un entorno económico incierto.

## 17. Conclusiones

### 17.1 Evaluación y Respuestas los Objetivos Planteados

**¿Cómo se ha desarrollado el ecosistema empresarial español en función del PIB e IPC, y qué implicaciones tienen estos indicadores en la estructura empresarial del país?**

El análisis de la influencia del Producto Interno Bruto (PIB) y del Índice de Precios al Consumidor (IPC) sobre las disoluciones y constituciones empresariales en España mediante modelos de regresión lineal múltiple muestra que ambos indicadores tienen un impacto limitado. Para las disoluciones empresariales, los resultados indican un coeficiente de determinación  $R^2$  de solo 0.001 y un  $R^2$  ajustado de -0.009, con un F-estadístico de 0.08673 y un p-valor de 0.917, lo que señala una relación no significativa entre estas variables macroeconómicas y el número de disoluciones. Los coeficientes para el IPC y el



PIB son insignificantes, con valores de -6.4564 y  $6.611 \times 10^{-10}$  respectivamente, y p-valores de 0.890 y 0.744, indicando una influencia casi nula sobre las disoluciones.

De forma similar, para las constituciones empresariales, el modelo revela un  $R^2$  de 0.003 y un  $R^2$  ajustado de -0.007, con coeficientes no significativos para el IPC y el PIB. Esto refuerza la observación de que las variaciones en el PIB y el IPC no están directamente correlacionadas con la tasa de nuevas empresas establecidas, sugiriendo que otros factores podrían jugar un rol más determinante en la dinámica empresarial del país.

### **¿Han afectado las variables macroeconómicas a la tasa de creación y disolución de empresas en España?**

Los modelos de regresión lineal múltiple ajustados para las tasas de constitución y disolución de empresas en España muestran que las variables macroeconómicas tienen una influencia moderada sobre estos procesos empresariales. Las variables explican un 6.0% de la variabilidad en las tasas de constitución, con un coeficiente ajustado R-cuadrado de 0.032, indicando una influencia moderada del entorno macroeconómico en la creación de empresas. Para las disoluciones, las variables macroeconómicas explican un 4.6% de la variabilidad, sugiriendo un impacto similarmente limitado en la tasa de disolución de empresas.

Aunque ciertos factores como el PIB y el IPC muestran efectos estadísticamente significativos, la capacidad de estos indicadores para explicar completamente las dinámicas de constitución y disolución es baja. Esto resalta la necesidad de considerar otros elementos contextuales y sectoriales para desarrollar políticas económicas efectivas. La investigación enfatiza que, para fomentar un entorno empresarial robusto y resiliente, es crucial adoptar un enfoque más holístico que integre diversas variables económicas, contextuales y sectoriales.

### **¿Qué variables específicas influyen en la constitución y disolución de empresas?**

El estudio sobre la influencia de variables en la constitución y disolución de empresas en España destaca la importancia del Indicador de Clima Empresarial (ICEA), significativo en ambos modelos de regresión. Para las constituciones empresariales, el ICEA presentó un coeficiente positivo significativo (coef = 38.323,  $p = 0.001$ ), sugiriendo que una mejor percepción del clima empresarial fomenta la creación de nuevas empresas. Este efecto positivo refleja cómo la confianza económica puede motivar a los emprendedores a iniciar nuevas empresas bajo expectativas de condiciones de mercado favorables y apoyo efectivo.

En contraste, para las disoluciones empresariales, el ICEA también fue significativo (coef = 8.3006,  $p = 0.003$ ) pero con un coeficiente menor, indicando que un ambiente empresarial favorable disminuye la probabilidad de disolución de empresas. Esto implica que en un entorno empresarial positivo, las empresas existentes tienden a sobrevivir y prosperar, enfrentando menos riesgos de cierre.

El papel del ICEA en ambos procesos resalta cómo un clima empresarial robusto no solo atrae nuevos emprendimientos sino también ayuda a la estabilidad y expansión de las empresas existentes. Aunque se consideraron otras variables económicas como el IPC y el PIB, estas no mostraron un impacto significativo, destacando la relevancia del ICEA como un indicador crítico que refleja factores económicos complejos y su influencia en el ecosistema empresarial español.

### **¿Cómo han respondido las empresas españolas ante las incertidumbres generadas por crisis macroeconómicas y cuáles son los factores clave para su resiliencia?**

Las empresas españolas han respondido a las crisis macroeconómicas con diversas estrategias que reflejan su capacidad de adaptación y resiliencia. Un factor clave identificado es la flexibilidad operativa, que permite a las empresas ajustar rápidamente sus operaciones y recursos frente a cambios económicos repentinos. Además, la diversificación de mercados y fuentes de ingresos ha sido crucial, permitiendo a las empresas no depender exclusivamente de un mercado o sector económico particular. Durante la entrevista con Alberto de Santos, se mencionó la importancia de la inteligencia de negocio y el análisis predictivo. Empresas que invierten en tecnologías de análisis avanzado y aprendizaje automático para entender mejor el entorno económico y prever cambios, tienden a manejar mejor la incertidumbre y a tomar decisiones más informadas. Estas herramientas permiten anticipar tendencias y adaptar estrategias en tiempo real, lo que es fundamental para mantener la competitividad y la sostenibilidad en tiempos de incertidumbre económica.

## **18. Opiniones Profesionales**

### **Relación con el Trabajo de Fin de Grado**

Ambos entrevistados perspectivas valiosas y complementarias a la investigación. Alberto, con su experiencia como emprendedor y en analítica de negocios en grandes corporaciones, destaca la relevancia de los factores macroeconómicos en la toma de decisiones empresariales. Por otro lado, José María, mediante su liderazgo en Controlnet y su habilidad para expandir operaciones en mercados internacionales, enfatiza la necesidad de adaptarse a condiciones macroeconómicas y regulatorias para alcanzar el éxito

empresarial. Juntos, estos perfiles enriquecen el TFG al ofrecer un marco robusto para comprender cómo el entorno económico afecta las decisiones estratégicas empresariales, resaltando la interacción entre variables macroeconómicas y dinámicas empresariales.

## **18.1 Resumen de la Respuestas Obtenidas**

### ***Consideración de la Situación Macroeconómica de Alberto y su equipo en el Desarrollo de Koulu (Startup de educación)***

Alberto explicó que durante el desarrollo inicial de Koulu, se tuvo en cuenta la situación macroeconómica de la época, específicamente en torno a la crisis inmobiliaria y la incertidumbre económica que persistía en esos años. El equipo de Koulu realizó un estudio para entender el gasto promedio de las familias en educación adicional, un factor crucial dado que Koulu estaba destinado a ser financiado por los padres de familia. Este análisis se apoyó en investigaciones previas de firmas multinacionales para establecer un marco de referencia sobre los gastos educativos en relación con el ingreso medio en España.

### ***Interés de Alberto en la Situación Macroeconómica durante las Estancias en Hewlett-Packard, Vodafone y Levi***

En esta pregunta Alberto destacó la importancia del contexto macroeconómico tanto en su interés personal como profesional. En su experiencia, el desempeño macroeconómico influye directamente en cosas como la evaluación del desempeño laboral a través del rendimiento económico que tiene la empresa donde se trabaja.

Alberto también contaba, como en Levi, específicamente, llegó a realizar un estudio para comprender cómo la inflación impactaba en los productos sustitutivos ofrecidos por la empresa y el patrón de compra de los clientes. Este estudio se vio interrumpido por el inicio del conflicto entre Rusia y Ucrania, que desestabilizó las tendencias económicas existentes. Así mismo, enfatizó que el interés en la macroeconomía puede que haya evolucionado de una mera curiosidad a una preocupación esencial para la supervivencia profesional, especialmente en el contexto de las recientes inestabilidades económicas globales.

### ***Consideración de la Situación Macroeconómica de José María y su equipo en el desarrollo y lanzamiento de Controlnet (Empresa de Desarrollo de software personalizado)***

En cuanto a si se tuvo en cuenta la situación macroeconómica al lanzar el producto, José María Martín-Mateos señaló que, durante la época de lanzamiento, el sector de las telecomunicaciones y soluciones empresariales digitales estaba en auge. Aunque la situación de capital de las empresas era relevante debido a la crisis inmobiliaria, la situación

macroeconómica general no fue un factor decisivo para el lanzamiento al mercado. Sin embargo, sí se observó y estudió el crecimiento en la venta de productos tecnológicos en otras empresas, que mostraba un aumento exponencial.

***Importancia de las situaciones o variables macroeconómicas del momento para la toma de decisiones en Controlnet.***

Respecto a la importancia de las variables macroeconómicas en la toma de decisiones para entrar en algunos países, destacó que ciertos indicadores, como el índice Big Mac o las políticas jurídicas de un país, son considerados para evaluar cómo podrían afectar los proyectos a realizar. Además, Martín-Mateos sugiere analizar indicadores económicos de líderes en los sectores de telecomunicaciones y digitalización, como Estados Unidos o Alemania.

*Para una visión completa de las entrevistas véase el Anexo 3*

### **18.3 Insights obtenidos de las preguntas**

Este proyecto investigó las estrategias empresariales influenciadas por variables macroeconómicas, combinando entrevistas con expertos, análisis de datos, y revisión literaria. José María compartió cómo, en el contexto de Controlnet, el auge tecnológico y aspectos como el índice Big Mac y las regulaciones jurídicas influenciaron sus decisiones de mercado, destacando el apoyo crucial de las administraciones públicas en tiempos de crisis. Alberto discutió la relevancia del análisis macroeconómico en multinacionales y startups como Koulu, enfatizando la adaptación a cambios económicos y la inclusión de variables no convencionales para una comprensión más detallada de las dinámicas empresariales. Ambos expertos subrayaron la necesidad de un enfoque multidimensional que combine condiciones macroeconómicas y características sectoriales, proveyendo una base robusta para entender la interacción entre economía y empresa, útil para futuras investigaciones y estrategias empresariales.

## 19. Recomendaciones y Propuestas

### 19.1 Recomendaciones

En el apartado de recomendaciones de este trabajo, se abordan estrategias y consejos dirigidos tanto a empresarios y emprendedores como a formuladores de políticas, con el objetivo de mejorar sus capacidades de adaptación y toma de decisiones en un entorno económico cambiante. Este segmento del estudio se estructura en dos subsecciones principales: la primera enfocada en proporcionar orientación práctica a empresarios y emprendedores, resaltando la importancia de integrar un análisis macroeconómico sólido en sus estrategias empresariales; y la segunda destinada a los formuladores de políticas, sugiriendo maneras en las que pueden apoyar y mejorar el ecosistema empresarial a través de políticas informadas y efectivas. A través de estas recomendaciones, se busca no solo aportar al debate académico, sino también ofrecer herramientas aplicables que puedan tener un impacto real y positivo en el entorno empresarial.

#### 19.1.1 Empresarios, Ejecutivos y Emprendedores

A lo largo de este estudio, el análisis de las variables macroeconómicas y su influencia en las dinámicas de constitución y disolución de empresas en España ha revelado aspectos clave que pueden ser cruciales para empresarios y emprendedores. A través de técnicas como la regresión lineal múltiple y el manejo de la multicolinealidad mediante el análisis de VIF, se ha identificado que, aunque variables como el PIB y la Deuda Pública fueron excluidas del modelo final por su alta correlación con otras variables, indicadores como el Indicador de Clima Empresarial (ICEA) tienen un impacto directo y significativo en el número de disoluciones empresariales. Este hallazgo destaca la importancia del clima empresarial y la confianza económica, subrayando cómo estos factores pueden influir directamente en la estabilidad y las decisiones estratégicas dentro del ecosistema empresarial.

A través de las entrevistas realizadas a profesionales del sector, se ha recopilado una visión que complementa los hallazgos técnicos de la investigación. Estas conversaciones han puesto de manifiesto la relevancia de tener una comprensión profunda del entorno económico cuando se toman decisiones estratégicas. En el contexto del desarrollo de productos y servicios, por ejemplo, se observó que un entendimiento detallado de la situación macroeconómica del sector en el que se opera puede ser decisivo. Esta perspectiva es vital ya que orienta a los emprendedores a considerar no solo el estado

actual del mercado, sino también a anticipar posibles escenarios económicos que podrían afectar sus modelos de negocio.

La investigación sugiere que para los empresarios y emprendedores, especialmente aquellos en las fases iniciales de desarrollo de un proyecto o negocio, es esencial integrar un análisis económico riguroso en la planificación y estrategia empresarial. Esto no solo incluye la observación de indicadores económicos estándar como el PIB o el IPC, sino también indicadores más específicos del clima empresarial y la confianza económica como el ICEA o incluso variables directamente relacionadas con sus sectores. Estos factores pueden proporcionar señales tempranas sobre las tendencias de mercado y potenciales desafíos o oportunidades que podrían surgir.

Además, la capacidad para adaptar y afinar continuamente los modelos predictivos y analíticos en respuesta a nuevos datos y tendencias emergentes es una competencia que puede diferenciar a las empresas exitosas de aquellas que no logran ajustarse a las realidades del mercado. Esto implica no solo la recolección y análisis de datos, sino también una constante revisión y mejora de los modelos empleados, asegurando así que las estrategias empresariales estén bien informadas y sean robustas ante cambios económicos y sociales.

En resumen, este trabajo destaca la importancia de una integración efectiva de análisis macroeconómico en la toma de decisiones empresariales, enfatizando que la comprensión y adaptación a las condiciones económicas no solo es una práctica deseable, sino una necesidad imperativa en el dinámico entorno de negocios de hoy.

### **19.1.2 Formuladores de Políticas**

En el ámbito de las políticas públicas y la administración, el papel de las autoridades en la configuración de un entorno empresarial sólido y adaptable es fundamental. A partir de los hallazgos de este trabajo, donde se ha evaluado el impacto de diversas variables macroeconómicas en la dinámica empresarial, es evidente que indicadores como el Indicador de Clima Empresarial (ICEA) son esenciales para que los empresarios evalúen las situaciones económicas actuales y futuras. Por tanto, es crucial que las políticas públicas promuevan la comprensión y la precisión de este y otros indicadores económicos, para que sean herramientas útiles en la toma de decisiones empresariales.

Para mejorar el sistema empresarial desde la perspectiva de las políticas públicas, se recomienda enfocarse en la promoción y educación continua sobre la importancia de la macroeconomía en las estrategias empresariales. Esto incluye asegurarse de que los indicadores como el ICEA sean accesibles, comprensibles y ampliamente difundidos entre la

comunidad empresarial. Al mejorar la calidad y la accesibilidad de la información económica, se facilita a los empresarios la tarea de hacer proyecciones más precisas y de ajustar sus estrategias a las condiciones del mercado.

Además, dada la relevancia del conocimiento económico evidenciada en las entrevistas a profesionales del sector, se sugiere que las administraciones públicas fomenten la implementación de la macroeconomía como pilar fundamental en la educación. Este enfoque no solo debería limitarse a los programas de estudios empresariales o económicos, sino expandirse a diversas disciplinas, preparando a los estudiantes para comprender mejor cómo las fluctuaciones económicas pueden afectar todos los aspectos de la vida profesional y personal.

En resumen, las recomendaciones hacia los formuladores de políticas se centran en fortalecer la base educativa en economía para todos los estudiantes, promover la transparencia y el acceso a datos económicos fiables, y asegurar que los indicadores económicos como el ICEA sean comprendidos y utilizados efectivamente dentro del panorama empresarial. Estas medidas no solo beneficiarán a los empresarios actuales y futuros, sino que también contribuirán a la creación de un entorno empresarial más resiliente y adaptativo frente a los desafíos económicos globales.

## **19.2 Propuestas**

En esta sección del Trabajo de Fin de Grado, se presenta una propuesta de proyecto de negocio que busca aplicar prácticamente los hallazgos de la investigación sobre la influencia de variables macroeconómicas en el ecosistema empresarial español. El proyecto se fundamenta en desarrollar una plataforma que utiliza análisis de datos avanzados para proporcionar herramientas estratégicas tanto a inversores como a empresas. La intención es mejorar la toma de decisiones mediante el acceso a análisis predictivos y descriptivos, permitiendo a las empresas mejorar su planificación y estrategia operativa en un entorno económico volátil. Este enfoque busca no solo aplicar conocimientos académicos, sino también ofrecer soluciones prácticas que tengan un impacto tangible en el mercado, facilitando así una mejor comprensión y gestión de los riesgos macroeconómicos.

### **19.2.1 Tipo de Propuesta**

En el desarrollo de este Trabajo de Fin de Grado, no solo se buscaba aportar una investigación técnica y teórica, sino también aplicar y poner en práctica las competencias y habilidades de negocio adquiridas durante la carrera en Business Analytics. Este enfoque

dual refleja un compromiso no solo con el entendimiento académico, sino también con la aplicación práctica y concreta de este conocimiento en escenarios del mundo real.

Es con este objetivo en mente que se propone desarrollar un modelo de negocio innovador: un “marketplace” dirigido tanto a inversores como a empresas. Esta plataforma se fundamenta en el análisis robusto de datos, integrando los insights obtenidos a lo largo de la investigación del TFG. La propuesta se centra en proporcionar un espacio donde los inversores puedan acceder a análisis predictivos y descriptivos que les permitan tomar decisiones de inversión más informadas, mientras que las empresas pueden utilizar estos datos para mejorar su posicionamiento, estrategias de crecimiento y toma de decisiones operativas.

Este tipo de propuesta no solo capitaliza los hallazgos de la investigación, sino que también sirve como un puente entre la teoría y la práctica, ilustrando cómo las habilidades de análisis de datos y comprensión de las dinámicas macroeconómicas pueden tener aplicaciones directas en el ámbito empresarial.

### 19.2.2 Problema y Solución de la Propuesta

La propuesta desarrollada en este Trabajo de Fin de Grado aborda un **problema** central en el ámbito de las inversiones y la gestión empresarial: **la dificultad para comprender y anticipar el impacto de las fluctuaciones macroeconómicas en las finanzas empresariales**. Los inversores, especialmente los “angel investors” y otros actores del mercado privado, enfrentan el desafío constante de identificar oportunidades de inversión que equilibren adecuadamente el riesgo y el retorno en un contexto económico que es inherentemente volátil y complejo. Paralelamente, las empresas necesitan navegar por estas mismas aguas turbulentas buscando capital y asesoramiento estratégico que les permita adaptarse y florecer ante cambios económicos impredecibles. Frecuentemente, tanto inversores como empresas carecen de acceso a herramientas analíticas avanzadas y a la experiencia necesaria para realizar evaluaciones financieras y estratégicas profundas.

La **solución** propuesta consiste en el desarrollo de una **plataforma integral que utiliza tecnologías de análisis predictivo avanzado para vincular directamente las variables macroeconómicas con las cuentas de resultados de las empresas**. Esta herramienta innovadora proporcionará a los inversores análisis y datos precisos, facilitando decisiones de inversión más informadas. Adicionalmente, la plataforma servirá como un medio para que los inversores ofrezcan asesoramiento directo a las empresas en las que inviertan sobre cómo gestionar y planificar frente a posibles cambios en el escenario



económico. Este enfoque no solo optimiza las decisiones de inversión, sino que también fortalece a las empresas, haciéndolas más resilientes y estratégicamente ágiles.

Por lo tanto, esta propuesta no solamente facilita un método reactivo para la inversión y la planificación empresarial, sino que establece una base sólida para el asesoramiento y la toma de decisiones que beneficia a todos los actores implicados en el ecosistema de inversión y gestión empresarial. En resumen, esta plataforma propone una sinergia entre tecnología y análisis financiero que ayuda a mitigar los riesgos derivados de las incertidumbres macroeconómicas, beneficiando tanto a inversores como a empresas en el mercado.

### **19.2.3 Explicación en Profundidad**

#### **Introducción al Modelo de Negocio**

**HELTUP** constituye una avanzada plataforma de análisis predictivo, específicamente diseñada para relacionar distintas variables con indicadores financieros, proporcionando a inversores y compañías capacidades analíticas para prever y adaptarse a las fluctuaciones económicas globales que podrían impactar sus operaciones y resultados financieros.

#### **Componentes Clave del Modelo**

*Tecnología y Análisis de Datos:* HELTUP integra inteligencia artificial y técnicas de aprendizaje automático para procesar y analizar grandes volúmenes de datos económicos y financieros. Esto facilitará la predicción del impacto de cambios macro en la actividad empresarial, apoyando la optimización de decisiones de inversión y estrategias empresariales.

*Interfaz y Experiencia del Usuario:* La plataforma se destaca por una interfaz intuitiva que incluye dashboards personalizables y reportes exhaustivos, permitiendo a los usuarios acceder fácilmente a evaluaciones de riesgo, proyecciones de mercado y recibir recomendaciones ajustadas a sus perfiles y necesidades.

*Integración y Seguridad de Datos:* La seguridad es primordial en el proyecto, asegurando la integridad y confidencialidad de los datos a través de avanzados protocolos de seguridad, que incluyen cifrado, autenticación robusta y medidas proactivas para la protección continua de la información de los usuarios.

#### **Funcionalidades y Servicios Ofrecidos**

HELTUP ofrece una serie de servicios críticos como análisis de riesgo, proyecciones de mercado y recomendaciones basadas en modelos predictivos sofisticados. Estos

servicios son vitales tanto para inversores que desean maximizar rendimientos como para empresas que buscan fortalecer su planificación estratégica ante cambios económicos inesperados. Estas herramientas benefician desde empresas en búsqueda de capital hasta inversores privados y analistas financieros.

### **Modelo de Monetización**

La monetización de la plataforma se realiza a través de modelos de suscripción, ofreciendo diferentes niveles de acceso adaptados a las necesidades específicas de los usuarios. Además, la plataforma proporciona informes especializados y servicios de consultoría a un costo adicional, así como licencias para grandes corporaciones que deseen integrar estas herramientas en sus sistemas internos.

### **Estrategias de Implementación y Escalabilidad**

Inicialmente, el negocio se lanzará en una fase de prueba piloto con un grupo selecto de usuarios para refinar funcionalidades y validar la robustez del sistema. La captación de usuarios iniciales se realizará mediante estrategias de marketing digital y colaboraciones con redes de inversores y entidades empresariales.

## **20. Resumen Narrativo**

Este trabajo de Fin de Grado ha explorado la influencia de las variables macroeconómicas en el ecosistema empresarial español, extendiéndose más allá del impacto inicial de la pandemia para abordar un espectro más amplio de factores económicos que afectan tanto la constitución como la disolución de empresas. El análisis ha evolucionado desde un enfoque inicial centrado exclusivamente en la pandemia, hacia un estudio más amplio que considera una variedad de variables macroeconómicas, adaptándose a las complejidades y a la falta de datos específicos sobre el COVID-19. A través de un proceso iterativo de ajuste metodológico, se han reformulado las preguntas de investigación para capturar mejor la dinámica del entorno empresarial en relación con el PIB, el IPC, y especialmente el Indicador de Clima Empresarial (ICEA), que ha demostrado ser significativo en el análisis.

Además, se ha integrado la perspectiva de expertos a través de entrevistas, complementando los análisis cuantitativos con experiencias y observaciones cualitativas que enriquecen los hallazgos. Este enfoque mixto ha permitido no solo validar los modelos estadísticos utilizados, sino también profundizar en la comprensión de cómo las empresas españolas se adaptan y responden a las incertidumbres macroeconómicas. La investigación concluye con una serie de recomendaciones prácticas para empresarios y formuladores de

políticas, enfocadas en fortalecer la resiliencia y el crecimiento del ecosistema empresarial frente a futuras crisis económicas.

Este estudio no solo contribuye al conocimiento académico, sino que también propone soluciones prácticas, destacando la importancia de la adaptabilidad y la previsión en la gestión empresarial y la formulación de políticas. Así, se establece un puente entre la teoría y la práctica, subrayando cómo los insights obtenidos pueden aplicarse para mejorar el entorno empresarial y económico en España.

## 21. Bibliografía

Cámara de Comercio de España. (n.d.). La creación de empresas. Recuperado de <https://www.camara.es/publicaciones/creacionempresas>

Análisis de los factores determinantes de la creación de empresas: una evidencia empírica en Castilla y León. (n.d.). Recuperado de <https://dialnet.unirioja.es/download/articulo>

Carr, R. (1966). Spain, 1808-1939. Oxford: Oxford University Press.

Nadal, J. (1975). El fracaso de la revolución industrial en España, 1814-1913. Barcelona: Ariel.

Ringrose, D. R. (1996). Spain, Europe, and the "Spanish Miracle", 1700-1900. Cambridge: Cambridge University Press.

Tortella, G. (1994). Los orígenes del capitalismo en España: banca, industria y ferrocarriles en el siglo XIX. Madrid: Tecnos.

Blanco, R., Mayordomo, S., Menéndez Pujadas, Á., & Mulino Ríos, M. (2021, August 24). El Impacto de la crisis del covid-19 sobre la vulnerabilidad Financiera de las empresas españolas. Repositorio Institucional: Análisis económico e investigación. <https://repositorio.bde.es/handle/123456789/17431>

Pérez-Calle, Mg. R. D., García-Casarejos, Dra. N., & García-Bernal, Dr. J. (2021, septiembre). La empresa española ante la COVID-19: factores de adaptación al nuevo escenario. RETOS. Revista de Ciencias de la Administración y Economía. [http://scielo.senescyt.gob.ec/scielo.php?script=sci\\_arttext&pid=S1390-86182021000100005](http://scielo.senescyt.gob.ec/scielo.php?script=sci_arttext&pid=S1390-86182021000100005)

Carr, R., & Fusi, J. P. (1987). Spain: Dictatorship to Democracy. London: Routledge.

Malefakis, E. E. (1970). Agrarian Reform and Peasant Revolution in Spain: Origins of the Civil War. New Haven: Yale University Press.

Payne, S. G. (1993). Franco and Hitler: Spain, Germany, and World War II. New Haven: Yale University Press.

Preston, P. (1996). The Spanish Civil War: Reaction, Revolution, and Revenge. New York: W. W. Norton & Company.

Viñas, Á. (1984). Guerra, dinero, dictadura: Ayuda fascista y autarquía en la España de Franco. Barcelona: Crítica.

Blanco, R., & Mayordomo, S. (2023, agosto 24). Evidencia sobre el alcance de los programas de garantías públicas y de Ayudas DIRECTAS a las empresas españolas implementados Durante La crisis del covid-19. SSRN. [https://papers.ssrn.com/sol3/papers.cfm?abstract\\_id=4540135](https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4540135)

Castelló, V. (2021, abril 12). El Impacto de la crisis en las empresas españolas tras un año de pandemia. El País. <https://elpais.com/economia/estar-donde-estes/2021-04-12/el-impacto-de-la-tesis-en-las-empresas-espanolas-tras-un-ano-de-pandemia.html>

Gutiérrez, H. (2023, febrero 10). El Cierre de Empresas en España en 2022 batió récords: 26.207 disoluciones, un 10% más. El País. <https://elpais.com/economia/2023-02-10/el-cierre-de-empresas-en-espana-en-2022-batio-record-26207-disoluciones-un-10-mas.html>

Guillén, M. F. (1997). The Limits of Convergence: Globalization and Organizational Change in Argentina, South Korea, and Spain. Princeton, NJ: Princeton University Press.

Holman, O. (1996). Integrating Southern Europe: EC Expansion and the Transnationalization of Spain. London: Routledge.

Dirección General de Economía y Estadística. (2021, December 2). La evolución económica y financiera de las empresas españolas en 2020 y 2021 según la central de balances. BANCO DE ESPAÑA (Eurosistema).

Banco de España. (s.f.). <https://www.bde.es/wbe/es/>

Molina, S., & Mora, A. (2010). La crisis financiera y las PYMES en España. Madrid: Fundación de las Cajas de Ahorros.

Ruesga, S. M., & Da Silva, A. (2011). La crisis económica y su impacto en las empresas españolas. Revista de Economía Aplicada, 19(57), 5-28.

Datos.gob.es. Gobierno de España (Conjuntos de Datos). (n.d.). [https://datos.gob.es/es/catalogo?administration\\_level=E&theme\\_id=sector-publico](https://datos.gob.es/es/catalogo?administration_level=E&theme_id=sector-publico)

# ANEXOS

## Anexo 1: Tablas

### Descripción de los datos

**Tabla 9.** Variables de LIMP.BBDD 3 clasificadas por tipo de variable e información sobre ellas

Nombre	Tipo de Dato	Descripción	Información Adicional
Column1	Integer	Número de observación, desde 0 hasta 627.	Índice de las filas.
Clase de disolución	String	Motivo de la disolución de las empresas.	Valores posibles: "Voluntaria", "Por fusión", "Otras".
Comunidad Autónoma	String	Nombre de las Comunidades Autónomas de España.	Incluye el código numérico de la comunidad.
Año	Integer	Año de la observación, desde 2012 hasta 2022.	
Número Sociedades	Integer	Ssociedades disueltas según clase, año y comunidad.	

**Tabla 10.** Variables de LIMP.BBDD 5 clasificadas por tipo de variable e información sobre ellas

Nombre	Tipo de Dato	Descripción	Información Adicional
Column1	Integer	Número de observación, desde 0 hasta 170.	Índice de las filas.
Indicador	String	Índice de confianza empresarial armonizado.	Solo toma el valor "ICEA".
Comunidad Autónoma	String	Nombre de las Comunidades Autónomas de España.	Incluye el código numérico de la comunidad.
Año	Integer	Año de la observación, desde 2013 hasta 2022.	
Valor Indicador	Integer	Valor del índice de confianza empresarial armonizado.	

NOTA: El ICEA (Índice de Confianza Empresarial Armonizado) es una medida que evalúa la confianza de las empresas en la economía y su capacidad para tomar decisiones de inversión.

**Tabla 11.** Variables de LIMP.BBDD 6 y 7 clasificadas por tipo de variable e información sobre ellas

Nombre	Tipo de Dato	Descripción	Información Adicional
<b>Column1</b>	Integer	Número de observación, desde 0 hasta 124.	Índice de las filas.
<b>Actividad económica</b>	String	Actividad económica según la CNAE-2009.	Incluye "Total Actividades"
<b>Comunidad Autónoma</b>	String	Nombre de las Comunidades Autónomas de España	"20 Nacional" representa el conjunto de España.
<b>Tipo de ERTE</b>	String	Tipo de ERTE.	Valores: "Total", "Parcial".
<b>Número de ERTES</b>	Integer	Número de ERTES según actividad, año y comunidad.	
<b>Año</b>	Integer	Año de la observación, desde 2013 hasta 2022.	

NOTA: Los ERTE (Expedientes de Regulación Temporal de Empleo) son medidas laborales que permiten a las empresas suspender temporalmente los contratos o reducir la jornada de empleados en situaciones excepcionales.

**Tabla 12.** Variables de LIMP.BBDD 8 clasificadas por tipo de variable e información sobre ellas

Nombre	Tipo de Dato	Descripción	Información Adicional
<b>Column1</b>	Integer	Número de observación, desde 0 hasta 170.	Índice de las filas.
<b>Fecha</b>	Date	Fecha de la observación en formato "dd/mm/aa".	Desde "01/03/2020" hasta "30/12/2021".
<b>Afiliados a la Seguridad Social</b>	Integer	Número de afiliados a la seguridad social según la fecha.	Datos diarios de la evolución de los ERTES.

## Medidas de Normalidad

**Tabla 13.** Medidas de Normalidad

Prueba	Variable	Estadístico	P-valor
<b>Shapiro-Wilk</b>	Población	0.80198586	2.858874868638953e-37
	Número de Sociedades	0.382824779	0.0
<b>Andersin-Darling</b>	Población	99.39626103	-
	Número de Sociedades	317.99002278	-
<b>D'Agostino's k^2</b>	Población	234.1741766	1.411638972451423e-51
	Número de Sociedades	1085.5180825	1.9175353385809348e-236

**Tabla 14. Medidas de Normalidad**

Prueba	Variable	Estadístico	P-valor
Shapiro-Wilk	Población	0.80198586	2.858874868638953e-37
	Capital Desembolsado	0.615095139	0.0
Andersin-Darling	Población	99.39626103	-
	Capital Desembolsado	196.81713157	-
D'Agostino's k^2	Población	234.1741766	1.411638972451423e-51
	Capital Desembolsado	597.45929103	1.8338481535740857e-130

## Medidas de Tendencia Central

**Tabla 15. Medidas de tendencia central de la segunda base de datos.**

Nombre	Media	Mediana	Moda
Año	2017	2017	2012
Número Sociedades	109,09	98,17	13,8

**Tabla 16. Medidas de tendencia central de la tercera base de datos.**

Nombre	Media	Mediana	Moda
Año	2017	2017	2012
Número Sociedades	397,04	124	0

Nombre	Media	Mediana	Moda
Año	2017,5	2017,5	2013
Valor Indicador	124,36	126,5	132

**Tabla 17. Medidas de tendencia central de la cuarta base de datos.**

**Tabla 18. Medidas de tendencia central de la quinta base de datos.**

Nombre	Media	Mediana	Moda
Año	2021,5	2021,5	2021
Número de ERTES	3934,27	1777	17

**Tabla 19. Medidas de tendencia central de la sexta y séptima base de datos.**

Nombre	Media	Mediana	Moda
Afiliados a la Seguridad Social	915199,72	705812	705812



## Medidas de Dispersión

**Tabla 20.** Medidas de dispersión de la segunda base de datos.

Nombre	Rango	Varianza	Desviación Estándar	Rango Intercuartílico
Año	10	10,04	3,17	6
Número Sociedades	280	5013,29	70,80	87,87

**Tabla 21.** Medidas de dispersión de la tercera base de datos.

Nombre	Rango	Varianza	Desviación Estándar	Rango Intercuartílico
Año	10	10,02	3,16	6
Número Sociedades	6 907	593934,52	770,67	350

**Tabla 22.** Medidas de dispersión de la cuarta base de datos.

Nombre	Rango	Varianza	Desviación Estándar	Rango Intercuartílico
Año	9	8,30	2,88	5
Valor Indicador	47	126,10	11,23	16,75

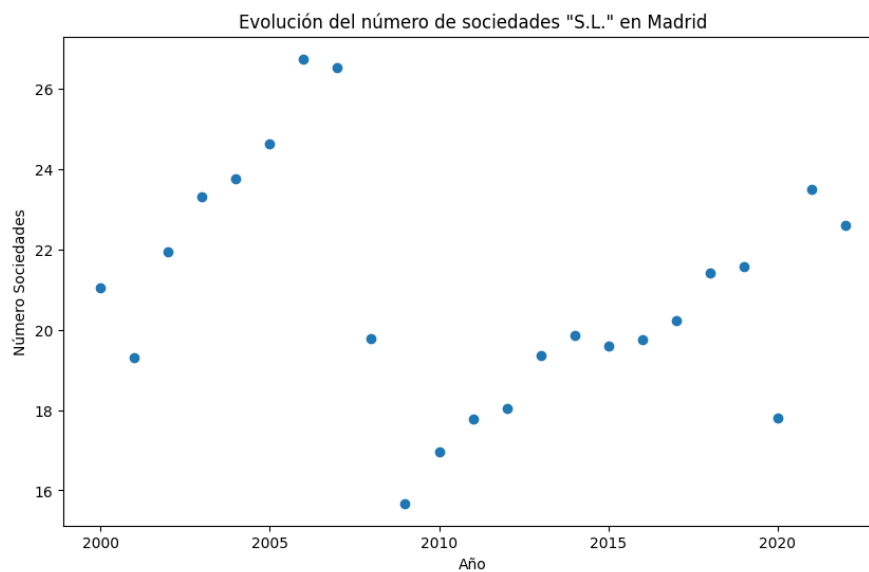
**Tabla 23.** Medidas de dispersión de la quinta base de datos.

Nombre	Rango	Varianza	Desviación Estándar	Rango Intercuartílico
Año	1	0,25	0,50	1
Número de ERTES	38829	31286261,52	5593,41	4189,5

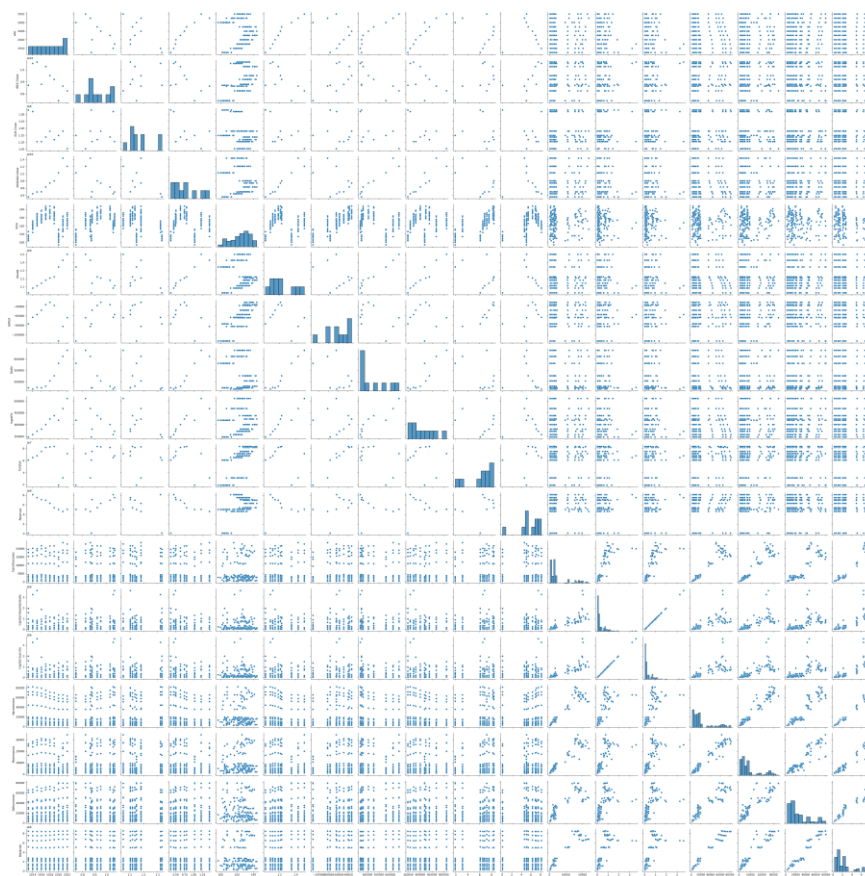
**Tabla 24.** Medidas de dispersión de la sexta base de datos.

Nombre	Rango	Varianza	Desviación Estándar	Rango Intercuartílico
Afiliados a la Seguridad Social	3616717	841433388713,49	917296,78	601390

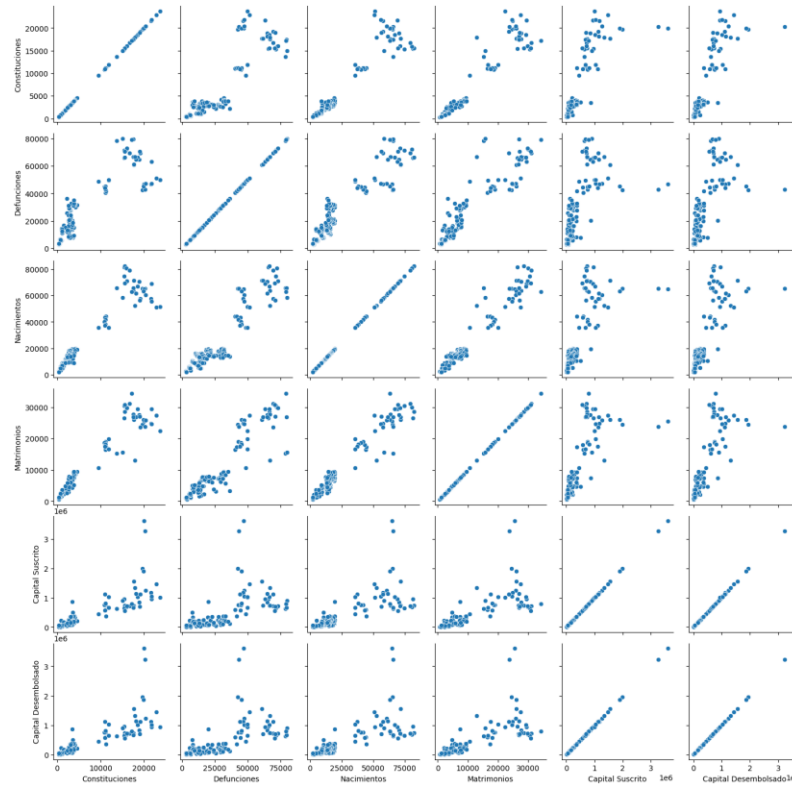
## Anexo 2: Ilustraciones



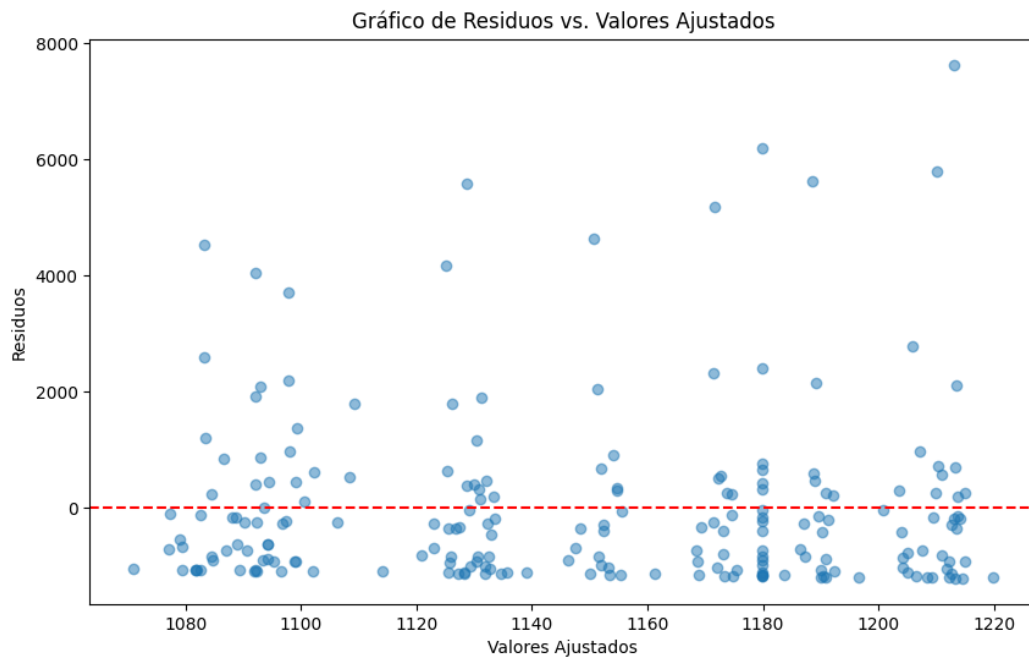
**Ilustración 9.** Captura de pantalla del diagrama de dispersión correspondiente a la evolución de sociedades de forma jurídica "S.L." constituidas en Madrid desde el año 2000 al 2022.



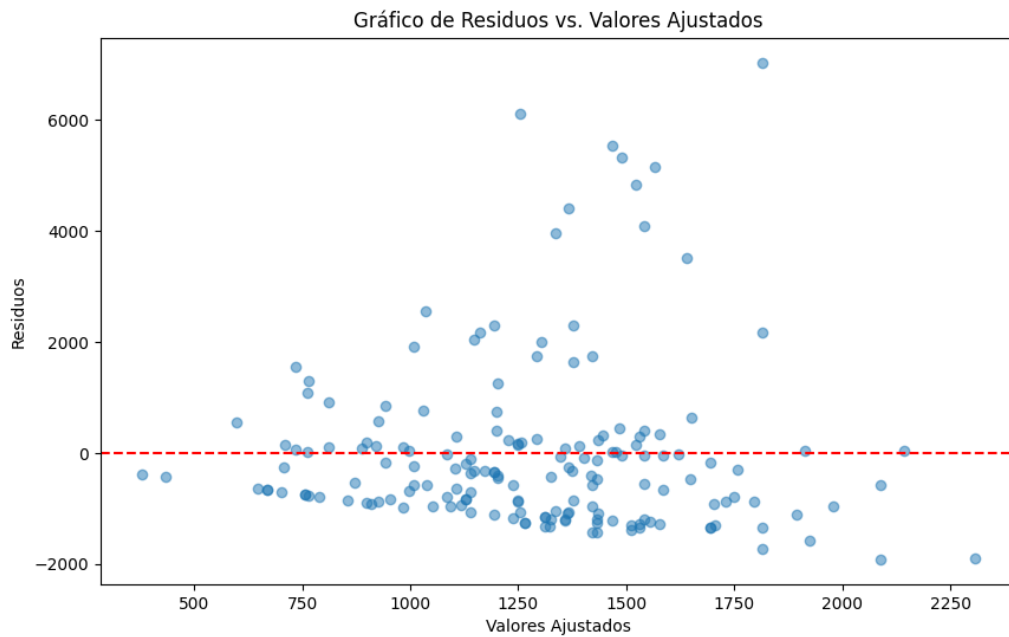
**Ilustración 10.** Gráfico de la matriz de dispersión para variables del modelo 3.



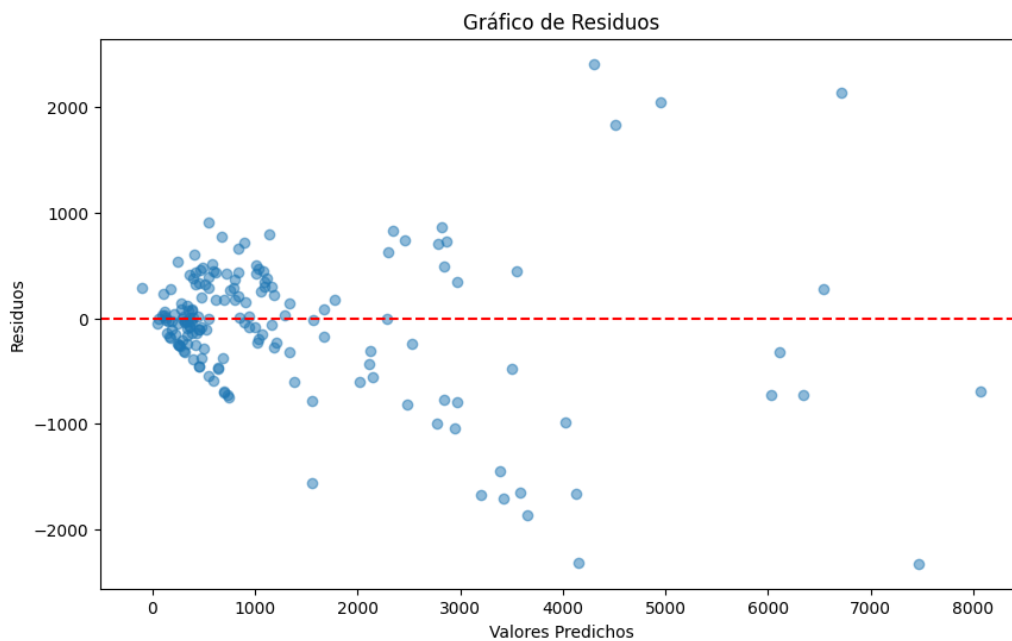
**Ilustración 11.** Gráfico de dispersión de variables específicas, en el contexto de constituciones empresariales.



**Ilustración 12.** Gráfico de residuos correspondiente al modelo1



**Ilustración 13.** Gráfico de residuos correspondiente al modelo2



**Ilustración 14.** Gráfico de residuos correspondiente al modelo3

## Anexo 3: Entrevistas Completas

### Información de las Entrevistas

#### Alberto de Santos (Perfil y Experiencia Profesional)

Alberto de Santos es un experto en analítica avanzada e inteligencia artificial, con más de quince años de experiencia en el sector, especializado en la toma de decisiones empresariales basadas en datos. Posee un profundo conocimiento en aprendizaje supervisado y no supervisado, así como en redes neuronales, y ha liderado equipos de científicos de datos en proyectos que maximizan el uso de Big Data e IA en negocios. Inició su carrera en Hewlett-Packard como Lead Data Scientist y co-fundó Koulu, una start-up de tutoría matemática con IA. Ha ocupado roles clave en Vodafone y actualmente es Global Senior Data Scientist en Levi Strauss & Co., compaginando con la docencia en inteligencia artificial en la Universidad Internacional de la Rioja y la Universidad Francisco de Vitoria.

#### José María Martín-Mateos Espinar (Perfil y Experiencia Profesional)

José María Martín-Mateos es un líder y emprendedor con más de veinte años de experiencia en tecnologías de la información y consultoría de gestión. Como fundador y CEO de Controlnet, parte del Grupo Lyvia, ha transformado la empresa en un líder en Sistemas de Gestión e Inteligencia de Negocio en Europa y Latinoamérica desde 2009. Antes de Controlnet, fue CEO de eresazul.com y COO en Grupo Martín-Mateos, donde consolidó su experiencia en gestión operativa y estratégica. Martín-Mateos se destaca por promover la innovación y adaptabilidad en un mercado tecnológico dinámico, con Controlnet ofreciendo soluciones en software, TI, marketing digital y consultoría creativa SaaS bajo la integración del grupo sueco LYVIA.

### Detalles de las Entrevistas

Las entrevistas realizadas a Alberto de Santos y José María Martín-Mateos, el 16 y 23 de abril de 2024 respectivamente, a través de plataformas de videollamada como Teams y Google Meets, proporcionaron insights significativos para el Trabajo de Fin de Grado sobre el impacto de las variables macroeconómicas en el entorno empresarial. Alberto discutió cómo estas variables influenciaron la fase de desarrollo de su startup y su percepción de su importancia en roles corporativos y educativos. Por su parte, José María se centró en cómo las condiciones macroeconómicas afectan las estrategias de desarrollo y sostenibilidad en empresas tecnológicas, cubriendo la evaluación de indicadores económicos y su impacto en la toma de decisiones empresariales.

## **Entrevista Alberto de Santos**

### ***Consideración de la Situación Macroeconómica de Alberto y su equipo en el Desarrollo de Koulu (Startup de educación)***

Alberto explicó que durante el desarrollo inicial de Koulu, se tuvo en cuenta la situación macroeconómica de la época, específicamente en torno a la crisis inmobiliaria y la incertidumbre económica que persistía en esos años. El equipo de Koulu realizó un estudio para entender el gasto promedio de las familias en educación adicional, un factor crucial dado que Koulu estaba destinado a ser financiado por los padres de familia. Este análisis se apoyó en investigaciones previas de firmas multinacionales para establecer un marco de referencia sobre los gastos educativos en relación con el ingreso medio en España.

### ***Interés de Alberto en la Situación Macroeconómica durante las Estancias en Hewlett-Packard, Vodafone y Levi***

En esta pregunta Alberto destacó la importancia del contexto macroeconómico tanto en su interés personal como profesional. En su experiencia, el desempeño macroeconómico influye y directamente en cosas como la evaluación del desempeño laboral a través del rendimiento económico que tiene la empresa donde se trabaja. Sin embargo, cree que actualmente y debido al contexto tan inestable que existe, puede llegar a afectar significativamente la estabilidad de los empleos dentro de las empresas.

Alberto también contaba, como en Levi, específicamente, llegó a realizar un estudio para comprender cómo la inflación impactaba en los productos sustitutivos ofrecidos por la empresa y el patrón de compra de los clientes. Este estudio se vio interrumpido por el inicio del conflicto entre Rusia y Ucrania, que desestabilizó las tendencias económicas existentes. Así mismo, enfatizó que el interés en la macroeconomía puede que haya evolucionado de una mera curiosidad a una preocupación esencial para la supervivencia profesional, especialmente en el contexto de las recientes inestabilidades económicas globales.

### ***Mejora del Análisis de Variables en el Modelo Analítico***

El entrevistado, considera que el porcentaje de explicación obtenido en el estudio, aunque pueda parecer modesto, es significativo en el contexto de las decisiones macroeconómicas y de negocios, donde incluso una pequeña fracción de predictibilidad puede ser valiosa. Subrayó que, en la vida real, las decisiones se toman con un conocimiento de aproximadamente un 40 o 50%, y tener un 5% o 6% de explicación en fenómenos complejos como las disoluciones y constituciones empresariales es un logro notable.

Sin embargo, sugiere que para mejorar el modelo y aumentar este porcentaje, sería beneficioso incorporar variables más diversificadas y ortogonales que no estén directamente relacionadas con los indicadores macroeconómicos habituales. Esto podría incluir factores como niveles educativos, tasas de desempleo locales, comportamientos de empresas específicas, inversiones en marketing, o incluso la situación económica de otros países que puedan influir en España.

### ***Atención de los Estudiantes a la Situación Macroeconómica***

De Santos compartió sus observaciones sobre el nivel de interés que sus estudiantes de inteligencia artificial muestran hacia las situaciones macroeconómicas, dependiendo del año de estudio en que se encuentran. Según explica, los estudiantes de primer y tercer año tienden a concentrarse más en sus estudios y en superar las exigencias académicas del momento, mostrando prácticamente ningún interés en los acontecimientos macroeconómicos. Sin embargo, observa un cambio en los estudiantes de cuarto año, quienes empiezan a considerar estos factores un poco más, aunque su enfoque principal sigue siendo encontrar empleo y ganar experiencia inmediata post-graduación. En general, percibe que la mayoría de los estudiantes no discuten ni consideran activamente los factores macroeconómicos en sus decisiones profesionales o académicas, enfocándose en lugar de eso en objetivos a corto plazo.

### ***Viabilidad y Valor de la Propuesta de Negocio***

Alberto evaluó la propuesta de negocio descrita para el TFG, calificándola de interesante y ambiciosa, particularmente en el contexto de un proyecto académico. Destacó la importancia y el desafío de articular y desarrollar una idea tan compleja en un entorno de trabajo de fin de grado. Desde una perspectiva profesional más práctica, sugirió reconsiderar el modelo de negocio para evitar la dependencia directa de las empresas en cuanto a la provisión de datos, identificándolo como un potencial riesgo. Propuso alternativas para obtener información relevante de forma independiente, como utilizar datos públicos de las empresas para realizar análisis y ofrecer servicios proactivos sin necesidad de solicitar datos adicionales directamente a las empresas. Su consejo se centra en transformar el enfoque hacia uno donde el valor se entregue de manera proactiva, posicionando el servicio como un recurso valioso y un socio de confianza para las empresas.

## **Entrevista José María Martín-Mateos**

### ***Consideración de la Situación Macroeconómica de José María y su equipo en el desarrollo y lanzamiento de Controlnet (Empresa de Desarrollo de software personalizado)***

En cuanto a si se tuvo en cuenta la situación macroeconómica al lanzar el producto, José María Martín-Mateos señaló que, durante la época de lanzamiento, el sector de las telecomunicaciones y soluciones empresariales digitales estaba en auge. Aunque la situación de capital de las empresas era relevante debido a la crisis inmobiliaria, la situación macroeconómica general no fue un factor decisivo para el lanzamiento al mercado. Sin embargo, sí se observó y estudió el crecimiento en la venta de productos tecnológicos en otras empresas, que mostraba un aumento exponencial.

### ***Importancia de las situaciones o variables macroeconómicas del momento para la toma de decisiones.***

Respecto a la importancia de las variables macroeconómicas en la toma de decisiones para entrar en algunos países, destacó que ciertos indicadores, como el índice Big Mac o las políticas jurídicas de un país, son considerados para evaluar cómo podrían afectar los proyectos a realizar. Además, Martín-Mateos tiende a analizar indicadores económicos de líderes en los sectores de telecomunicaciones y digitalización, como Estados Unidos o Alemania.

### ***Ayuda recibida por parte de las administraciones públicas durante las crisis macroeconómicas.***

Sobre el apoyo de las administraciones públicas durante crisis macroeconómicas, indicó que desde el principio, las administraciones públicas han apoyado a Controlnet no solo en visibilidad y difusión sino también en impulsar proyectos y acciones de beneficio mutuo. Por ejemplo, se mencionó el inicio de un ecosistema tecnológico en la provincia de Cádiz para fomentar el empleo y la formación profesional en el sector.

### ***En relación con las respuesta de Alberto de Santos, opinión sobre el enfoque de variables ortogonales y variables que sean interesantes estudiar y añadir al modelo?***

En cuanto al enfoque de utilizar variables ortogonales en el análisis, consideró que es una buena estrategia, aunque recomendó enfocarse en variables relevantes para el sector específico bajo estudio. Por ejemplo, en el sector de las telecomunicaciones, sugirió centrarse en las ventas anuales de servicios o productos tecnológicos de las principales



empresas de España o Europa, lo cual podría derivar en la creación de variables de ventas por cada sector importante para obtener una visión más clara del mercado.

***Conocimiento sobre el ICEA y recomendaciones de índices que utilice para seguir la macroeconomía***

Respecto al conocimiento y seguimiento de índices como el Índice de Confianza Empresarial Armonizado (ICEA), afirmó que no estaba familiarizado con este índice específicamente, pero utiliza índices similares en Alemania y EE. UU. para anticipar el comportamiento del sector. Considera que el seguimiento de estos índices puede proporcionar una perspectiva anticipada de las tendencias del mercado.

***Recomendaciones y consejos para emprendedores***

Finalmente, sobre las recomendaciones para nuevos emprendedores, destacó la importancia de tener un producto o servicio viable y escalable. Subrayó que la obsesión, el compromiso y el esfuerzo son esenciales para el éxito de un proyecto, lo que implica un enfoque total y la posibilidad de tener que sacrificar otros aspectos de la vida personal para construir un camino exitoso en el negocio.