



Facultad de Ciencias Económicas y Empresariales

Hipotecas y Desigualdad Regional en España: Desarrollo de un Modelo Analítico para Evaluar el Acceso a la Financiación

Clave: 202004334

ÍNDICE DE CONTENIDO

ÍNDICE DE ILUSTRACIONES.....	4
ÍNDICE DE TABLAS	5
RESUMEN	6
ABSTRACT.....	7
1. INTRODUCCIÓN	8
1.1. Objetivos del trabajo	9
1.2. Estructura del documento	10
2. ESTADO DEL ARTE.....	11
2.1. Factores macroeconómicos y evolución del mercado hipotecario.....	11
2.2. Políticas regulatorias	12
2.3. Desigualdad regional.....	12
2.4. Previsiones para el futuro	13
3. METODOLOGÍA Y HERRAMIENTAS UTILIZADAS.....	14
3.1. Descripción de la base de datos.....	14
3.2. Herramientas utilizadas para el análisis	15
3.2.1. Elección de Python	15
3.2.2. Elección de Google Colab	16
3.3. Visualización de los datos.....	16
4. DESARROLLO DEL SISTEMA.....	19
4.1. Interfaz de usuario.....	19
4.2. Funcionalidades principales	21
4.2.1. Comparación de comunidades autónomas	21
4.2.2. Generación de gráficos.....	25
4.2.3. Matriz de scores	27
4.3. Requisitos previos para la ejecución del código	27
5. RESULTADOS.....	28
5.1. Comparación de comunidades autónomas	28
5.2. Análisis de las gráficas	32
5.2.1. Gráficos de Dispersión	32
5.2.2. Mapa de calor de correlaciones	38

6. CONCLUSIONES.....	40
7. LIMITACIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN	42
Bibliografía.....	45
Anexo I: Código de programación	46
Anexo 2: Archivo Excel con “BBDD_TFG.xlsx”	53

ÍNDICE DE ILUSTRACIONES

Ilustración 1: Cálculo de Sedad	23
Ilustración 2: Cálculo de Score.....	23
Ilustración 3: Desplegables y botones de la interfaz	28
Ilustración 4: Ejemplo de perfil seleccionado	28
Ilustración 5: Ejemplo de comunidades y perfil a analizar	29
Ilustración 6: Ejemplo de comunidades y perfil a analizar	30
Ilustración 7: Ejemplo de comunidades y perfil a analizar	30
Ilustración 8: Matriz de scores para todas las CCAA.....	31
Ilustración 9: Relación entre PIB per cápita e Hipotecas per cápita	32
Ilustración 10: Relación entre PIB per cápita e Importe de hipotecas per cápita	33
Ilustración 11: Relación entre Sueldo Medio e Hipotecas per cápita.....	34
Ilustración 12: Relación entre Sueldo Medio e Importe de hipotecas per cápita	35
Ilustración 13: Relación entre Tasa de desempleo e Hipotecas per cápita	36
Ilustración 14: Relación entre Tasa de desempleo e Importe de hipotecas per cápita	37
Ilustración 15: Mapa de calor de correlaciones	38

ÍNDICE DE TABLAS

Tabla 1: Librerías utilizadas en el código	18
Tabla 2: Información relativa a los perfiles medios	20

RESUMEN

Este trabajo de fin de grado analiza los factores sociales, económicos y financieros que influyen en la concesión de hipotecas en España, enfocado en las diferencias que existen entre comunidades autónomas. En este estudio, se han considerado el PIB, la tasa de desempleo, el sueldo medio, la edad y el importe de las hipotecas como variables relevantes para entender su relación con la probabilidad de obtener una hipoteca. Se ha utilizado una metodología basada en la recopilación de datos estandarizados y gráficos de dispersión para identificar patrones y correlaciones entre las variables consideradas. Los resultados muestran que factores como el sueldo y la estabilidad laboral son determinantes en la concesión de hipotecas, mientras que la influencia que tienen la tasa de desempleo y el PIB per cápita es menos directa. Además, se observa que las políticas bancarias y las dinámicas regionales juegan un papel crucial, con comunidades como Madrid destacando por su mayor probabilidad de concesión de hipotecas. Este trabajo también expone las limitaciones del análisis, como la falta de datos cualitativos y la ausencia de variables clave como la inflación y los tipos de interés. Finalmente, se proponen futuras líneas de investigación, como la aplicación de técnicas de machine learning y el análisis a nivel provincial o internacional.

Palabras clave: hipotecas, desigualdades regionales, análisis de datos

ABSTRACT

This work analyzes the social, economic and financial factors that influence the granting of mortgages in Spain, focusing on the differences that exist between regions. In this study, GDP, unemployment rate, average salary, age and mortgage amount have been considered as relevant variables to understand their relationship with the probability of obtaining a mortgage. A methodology based on standardized data collection and scatter plots has been used to identify patterns and correlations between the variables considered. The results show that factors such as salary and job stability are determinants in the granting of mortgages, while the influence of the unemployment rate and GDP per capita is less direct. In addition, banking policies and regional dynamics are found to play a crucial role, with communities such as Madrid standing out for their higher probability of granting mortgages. This paper also exposes the limitations of the analysis, like the lack of qualitative data and the absence of key variables such as inflation and interest rates. Finally, future lines of research are proposed, such as the application of machine learning techniques and analysis at the provincial or international level.

Key words: mortgages, regional inequalities, data analysis

1. INTRODUCCIÓN

Este trabajo trata de estudiar los distintos escenarios de concesiones hipotecarias, a través de un simulador de hipotecas basado en datos socioeconómicos y financieros de cada comunidad autónoma. El simulador se propone como una herramienta analítica que permite evaluar la aprobación hipotecaria, permitiendo identificar y cuantificar posibles desigualdades entre comunidades autónomas en las concesiones hipotecarias.

A través de este trabajo, se busca contribuir al entendimiento de cómo las características económicas y sociales pueden llegar a influir en el acceso a la financiación hipotecaria. Este análisis no solo trata de mostrar la existencia de desigualdades en el sistema hipotecario español, sino también de abrir nuevas líneas de investigación y debate académico.

La desigualdad económica es un fenómeno que ha sido ampliamente estudiado en el ámbito tanto académico como político, pero el impacto que tiene en áreas específicas como el acceso a la financiación hipotecaria, sigue siendo un área de debate. España, es un país cuyas diferencias regionales son un factor determinante, el acceso a la vivienda representa uno de los mayores retos económicos y sociales. Este acceso está en muchas ocasiones condicionado por factores económicos, sociales y demográficos, lo que varía significativamente entre comunidades autónomas y muestra posibles disparidades estructurales.

Según (Soza Marañón, 2019), los créditos hipotecarios, conocidos como hipotecas, son el principal instrumento financiero para la adquisición de vivienda. Según (BBVA, BBVA, s.f.), una hipoteca es un préstamo bancario en el que el cliente recibe una cantidad de dinero a cambio del compromiso de devolverla con intereses mediante cuotas periódicas, con la particularidad de que el inmueble adquirido actúa como garantía. Esta característica vincula la concesión de hipotecas a la situación económica del solicitante y al contexto financiero de cada región

Además, noticias recientes afirman que el encarecimiento de la vivienda en España duplica el promedio de la Unión Europea, dificultando aún más la adquisición de inmuebles en ciertas regiones (SER, 2025). A pesar de este contexto: “La Unión de Créditos Inmobiliarios (UCI) estima que la concesión de hipotecas podría registrar un crecimiento del 10% durante este año, hasta alcanzar cerca de 400.000 operaciones, tras la reactivación en la concesión de hipotecas en los últimos meses de 2024 por el impulso

de la bajada de tipos de interés del Banco Central Europeo (BCE), lo que ha sentado las bases para un 2025 "prometedor" ” (Press, 2025).

Para este estudio, utilizaremos como base el análisis de la concesión de hipotecas como instrumento financiero clave para la adquisición de la vivienda. Las entidades financieras tienen un papel fundamental en este proceso ya que son las encargadas de evaluar a los distintos solicitantes a través de modelos de riesgo de crédito. Sin embargo, estos modelos pueden verse influenciados por características regionales.

1.1. Objetivos del trabajo

El principal objetivo de este trabajo es realizar un análisis exhaustivo sobre diferentes variables socioeconómicas que influyen en la concesión de hipotecas para así poder extraer conclusiones de cuáles son las más influyentes, si hay alguna relación entre ellas y, de alguna forma, predecir en qué comunidades autónomas es más fácil recibir una hipoteca en función del nivel socioeconómico del usuario.

Este estudio recoge información desde el año 2014 al año 2023, por lo que permite que el análisis sea temporal y así poder sacar conclusiones robustas evitando algún “outlier” que pudiera influir en el análisis.

A través de la generación de gráficos y el análisis de datos, el trabajo busca presentar los resultados de la manera más fácil y clara permitiendo a los usuarios, independientemente de su nivel de conocimiento sobre hipotecas, interpretar los datos de manera correcta.

Además, una parte muy importante del trabajo es proporcionar una herramienta lo más interactiva posible para que así el usuario, pueda usar la herramienta sin dificultad y pueda extraer sus propias conclusiones.

En conclusión, el objetivo principal de este trabajo es comprender cómo diversas variables socioeconómicas afectan el proceso de concesión de hipotecas en cada comunidad autónoma. Se busca analizar cómo factores como el nivel de ingresos, la tasa de desempleo, la edad promedio de los solicitantes, y otros aspectos demográficos y económicos influyen en la toma de decisiones de las entidades financieras. Para alcanzar este objetivo, se emplearán herramientas avanzadas de programación y análisis de datos, lo que permitirá procesar y visualizar grandes volúmenes de información. De este modo, se obtendrá una visión más detallada y clara de cómo funciona el mercado hipotecario en España, facilitando la identificación de patrones y tendencias que podrían ser útiles tanto para los agentes financieros como para los responsables de la política económica.

1.2. Estructura del documento

Este documento se estructura en 7 apartados de forma que lo que se quiere transmitir quede de la manera más estructurada y entendible posible.

El primer apartado consiste en una breve introducción donde se explicará el contexto y los objetivos principales del proyecto. Tras esta, el siguiente apartado consta del estado del arte, donde se analizarán estudios previos sobre el mercado hipotecario y sobre cómo influyen algunas variables en las diferentes comunidades autónomas.

El tercer apartado se centra en la metodología y las herramientas utilizadas a lo largo del desarrollo del proyecto. Tal como indica el título, en esta sección se detallarán y explicarán las distintas herramientas y técnicas empleadas, justificando su elección y describiendo cómo contribuyen a la correcta ejecución de cada fase del trabajo. Más adelante se encuentra el apartado de desarrollo del sistema en el que se describirá cómo funciona el código, cuáles son los diferentes modos que tiene y los requisitos necesarios para poder ejecutarlo.

Una vez establecido el marco metodológico y tecnológico, es el turno de observar y comentar los resultados con la interfaz y las gráficas obtenidas, este punto será el más importante del trabajo ya que será donde se explique en profundidad los objetivos del trabajo, observar si existe alguna relación entre las variables y ver las diferencias entre las comunidades autónomas.

El estudio concluirá con el apartado de las conclusiones, dedicado a la reflexión y a la síntesis de los resultados. Por último, el documento incluye una bibliografía con las fuentes consultadas para poder realizarlo y se adjuntará un anexo con el código de programación empleado.

2. ESTADO DEL ARTE

En este apartado vamos a ver cuáles son las principales aportaciones teóricas del tema, para ello estudiaremos algunos factores macroeconómicos que afectan directamente a esta temática. El acceso por parte de los ciudadanos a financiación hipotecaria ha sido siempre un tema muy relevante en la sociedad debido a su impacto en la economía. Existen varios factores que influyen en la concesión de hipotecas como las condiciones macroeconómicas, decisiones regulatorias de las entidades financieras, evolución del mercado inmobiliario y muchas otras condiciones.

El objetivo de este segundo punto es presentar una visión actual del conocimiento en torno a la concesión de hipotecas en España.

2.1. Factores macroeconómicos y evolución del mercado hipotecario

El mercado hipotecario español ha sufrido muchos cambios en estos últimos años debido a la evolución del PIB, la tasa de desempleo, la política de tipos de interés que se está llevando a cabo desde el Banco Central Europeo (BCE). Según un informe del (BBVA, Evolución reciente del mercado hipotecario, 2021), la concesión de hipotecas y la venta de inmuebles se ha recuperado del COVID-19, pero este crecimiento podría ser solamente temporal, debido a que puede deberse a acumulación de dinero ahorrado y sobre todo a condiciones de financiación ampliamente favorables para los ciudadanos. Otros motivos que podría indicar que es algo puramente temporal son el endurecimiento de las políticas monetarias y la inflación.

Por otra parte, el Banco Central Europeo ha estudiado las consecuencias de que los bancos restrinjan el acceso a hipotecas a los ciudadanos y se han dado cuenta que, si lo limitan, no solo se reduce la demanda para comprar inmuebles, sino que además hace que los precios de los alquileres se disparen, ya que habrá mucha más oferta para menos demanda porque más personas se verán obligadas a permanecer como inquilinos.

Estas consecuencias pueden ser muy graves para la economía, en este caso española, ya que podría aumentar la desigualdad económica, y una de las herramientas clave para generar patrimonio y riqueza son los bienes inmuebles.

2.2. Políticas regulatorias

En España, a partir de la crisis del 2008, las políticas regulatorias para la concesión de hipotecas se han endurecido a pasos agigantados. Por ejemplo, se estableció la Ley de Crédito Inmobiliario que “establece normas de protección y transparencia en los préstamos hipotecarios adquiridos por personas físicas”. (MontePrincipe Gestión, 2025). Esta ley se creó para establecer requisitos mucho más estrictos para la concesión de hipotecas y para obligar a los bancos a informar con mayor transparencia sobre los consumidores.

Según periódicos como El País y El Confidencial, han advertido que estas restricciones tan estrictas podrían estar dificultando el acceso a una vivienda a diversos grupos de la población como pueden ser los jóvenes y los trabajadores con empleos poco remunerados.

Además, al haber endurecido las concesiones de hipotecas, esta decisión está derivando en un incremento de propiedades en manos de inversionistas o grandes fondos, lo que hace que se limite aún más el acceso de familias a la compra de inmuebles.

2.3. Desigualdad regional

Uno de los aspectos más relevantes, y donde quiero centrar mi trabajo, es en la desigualdad regional que hay en España para acceder a financiación.

Hay diversos motivos que podrían afectar a esta disparidad, pero uno de los más importantes es el hecho de que la Ley de Crédito Inmobiliario se impuso a nivel nacional, lo que beneficia a comunidades con mayor actividad económica como Madrid y Cataluña y perjudica a comunidades con ciudadanos con niveles de ingresos menores y mayor desempleo como Extremadura o Castilla-La Mancha.

Además, en comunidades donde las tasas de desempleo son elevadas, hay un mayor riesgo de impago por lo que los bancos deben ser más reacios a conceder las hipotecas.

Para corroborar esta desigualdad regional, (20minutos, 2024) calculó el porcentaje de esfuerzo teórico anual que los hogares deben destinar al pago de la primera cuota de una hipoteca. Los resultados reflejan una brecha significativa: en Islas Baleares, la población local debe destinar el 64,4% de su renta disponible, seguida de Málaga (56,8%) y Barcelona (58,5%). En contraste, provincias como Castellón (23,2%), Teruel (23,4%) y

Ciudad Real (24%) presentan los porcentajes más bajos. España en su conjunto representa un esfuerzo medio de un 34,9%.

Como se puede observar, esto conlleva a un círculo vicioso en el que es muy difícil evolucionar, los ciudadanos con menor capacidad adquisitiva tienen menos acceso a la compra de inmuebles y además el precio de los alquileres aumentan, lo que reduce su capacidad de ahorro para poder acceder en el futuro a una posible compra.

2.4. Previsiones para el futuro

De entre las 10 tendencias del sector inmobiliario en 2025 detectadas por la UCI ((UCI), 10 tendencias del sector inmobiliario en 2025, 2025), la bajada de tipos de interés por parte del BCE, impulsará la concesión de hipotecas. Desde UCI estiman que en 2025 habrá un crecimiento de un 10%, lo que permitirá superar el umbral de las 400.000 operaciones.

A pesar de este impulso a las concesiones de hipotecas, siguen existiendo diferentes problemáticas, por ejemplo, el precio de la vivienda en algunas comunidades está aumentando de manera significativa, existe falta de oferta en muchas ciudades y sobre todo la inflación y las políticas regulatorias tendrán un papel muy importante en la concesión de las hipotecas.

3. METODOLOGÍA Y HERRAMIENTAS UTILIZADAS

3.1. Descripción de la base de datos

Para la realización de este trabajo, se ha creado una base de datos llamada *BBDD_TFG* que contiene información demográfica, económica y datos sobre hipotecas. Este archivo contiene información de todas las comunidades autónomas de España desde 2014 a 2023. El objetivo de crear esta base de datos es que el estudio que se detallará en puntos posteriores se sustente con datos fiables y de calidad.

Para obtener los datos que nutren este archivo se ha extraído la información del INE (Instituto Nacional de Estadística) y de los diferentes periódicos locales de cada comunidad autónoma, para que, de esta forma, la información quede totalmente contrastada. Para poder utilizar estos registros, la información se ha recopilado, juntado y estructurado para que puedan ser utilizados de manera eficiente en este estudio.

La base está organizada de forma que en cada columna están las diferentes variables que se han considerado relevantes para el estudio. Las variables están organizadas en 3 categorías:

- La primera categoría incluye las columnas de “*CCAA*” y “*Año*”, sirve como identificación. “*CCAA*” es el nombre de la comunidad autónoma a la que pertenecen los datos y “*Año*”, como el propio nombre indica es el año al que corresponden los datos de esa fila.
- La segunda categoría es la de datos demográficos y económicos. Las columnas son “*Población*” que es el número total de habitantes que la comunidad tenía ese año, “*Tasa de desempleo*” es el porcentaje de población que se encontraba en situación de desempleo, “*Tasa de desempleo de jóvenes*” y “*Tasa de desempleo de adultos*”, tienen exactamente la misma explicación que el punto anterior pero solamente contando los mayores y menores de 26 años respectivamente, “*Sueldo medio*” es la media de ingresos anuales brutos que tenían los trabajadores de esa comunidad autónoma y por último “*PIB*”, es el Producto Interior Bruto de la comunidad, ya que es una medida que se utiliza como indicador de la riqueza y progreso de cualquier país o región.
- La última categoría que se ha incluido es la de hipotecas. “*Número de hipotecas*” indica el número total de hipotecas que se concedieron, mientras que “*Importe de hipotecas*” indica la suma total del dinero prestado como hipoteca.

Además de estas columnas, cuyos datos son extraídos de internet se han creado 3 columnas extra, “*Hipotecas per cápita*” que indica el número de hipotecas que se concedieron por cada habitante, “*PIB per cápita*”, indica la riqueza media de cada habitante y “Importe de hipotecas per cápita” que indica el dinero total prestado en una hipoteca por cada habitante.

En la preparación de esta base de datos se han tenido en cuenta varios aspectos como la completitud de datos, homogeneidad temporal y unidades de medida para así garantizar la fiabilidad de este análisis.

3.2. Herramientas utilizadas para el análisis

Para el desarrollo de este Trabajo Fin de Grado se ha elegido Python como lenguaje de programación y Google Colab como plataforma de ejecución.

3.2.1. Elección de Python

Por los siguientes motivos se ha considerado a Python el mejor lenguaje de programación posible para ese proyecto:

- Facilidad de uso

La sintaxis de Python es sencilla y sobre todo intuitiva, por lo que eso hace que su escritura sea lo más fácil posible. (Vieira, HostGator, 2024)

- Amplia variedad de bibliotecas

Antes incluso de saber exactamente en qué iba a consistir este trabajo, se decidió que el programa que se utilizaría sería Python ya que te ofrece una amplia variedad de bibliotecas y en caso de tener que modificar el código y añadir alguna biblioteca extra no habría problema para encontrarla.

En este proyecto se han utilizado las bibliotecas “Panda” para el análisis y la manipulación de los datos provenientes de *BBDD_TDG*. También se han utilizado las bibliotecas “Matplotlib” y “Seaborn”, para la generación y visualización de las gráficas. Y por último “Ipywidgets”, que es una biblioteca que crea interfaces interactivas muy sencillas para su programación y para el entendimiento del usuario.

- Popularidad

El último motivo por el que se ha elegido este lenguaje de programación ha sido por la popularidad y el manejo que mucha gente ya tiene de esta herramienta. En caso de que alguien sin experiencia en programación necesite usar o revisar este código, Python es la opción más accesible, ya que cuenta con abundante documentación, tutoriales y foros de ayuda.

3.2.2. Elección de Google Colab

Google Colab es una plataforma ofrecida por Google. Proporciona un entorno de cuaderno Jupyter alojado en la nube que permite escribir y ejecutar código Python. (Vieira, HostGator, 2024)

Por los siguientes motivos se ha considerado a Google Colab como la mejor plataforma posible para ese proyecto:

- Sin instalación

No es necesario ni instalar Python, ni descargar las bibliotecas necesarias, simplemente se trabaja en línea.

- Facilidad para compartir

Al tratarse de un entorno en línea, se puede compartir el trabajo o editarlo desde cualquier ordenador.

- Integración con Google Drive

Para tener la base de datos, el código y la memoria almacenados en una misma carpeta, Google Drive te permite guardarlos directamente, lo que facilita su organización.

3.3. Visualización de los datos

Para entender mejor la relación entre las diferentes variables que ya se han definido en el apartado “3.1 Descripción de la base de datos”, se han creado diferentes gráficas a través de las bibliotecas “Matplotlib” y “Seaborn”.

Estas visualizaciones permiten identificar tendencias, patrones y correlaciones entre los datos disponibles.

Las principales gráficas utilizadas son las siguientes:

- Diagramas de dispersión

Es un diagrama que permite representar un conjunto de datos de dos variables en los dos ejes de las coordenadas cartesianas, es decir, sirven para analizar la relación entre dos variables estadísticas, se le puede llamar también diagrama de correlación. (Balderix, s.f.)

- Matriz de correlación y mapa de calor

La matriz de correlación es una tabla que muestra los coeficientes de correlación entre las diferentes variables elegidas. Cada celda indica la correlación entre dos variables, de esta forma, se consigue ver de una manera visual todas las relaciones.

Los valores de esta matriz oscilan entre -1 y 1, este último valor indica una correlación positiva perfecta, el -1 indica una correlación negativa perfecta y el 0 indica que no existe correlación entre las dos variables.

En este caso se ha elegido para calcular los coeficientes de correlación el método Pearson. A pesar de que la biblioteca de Python “Pandas” calcula la correlación de Pearson por defecto, también es considerada como la idónea para este estudio ya que se trata de analizar relaciones lineales entre las variables.

Además, la matriz de correlaciones se ha fusionado con un mapa de calor para que sea aún más visual, marcando en rojo una correlación positiva perfecta y en azul una correlación negativa perfecta. (LearnStatistics, s.f.)

Se han utilizado estos dos tipos de gráficos ya que se ha considerado que la mejor forma de visualizar y analizar la relación entre las variables elegidas son los gráficos de dispersión y la matriz de correlaciones.

A continuación, se muestra una tabla resumen con las librerías que han sido utilizadas en el código para la realización de este trabajo, con una breve descripción de la función que tiene cada una de ellas.

Tabla 1: Librerías utilizadas en el código

FUNCIÓN	LIBRERÍA
Lectura de archivos Excel	Pandas
Manipulación de arrays y matrices	Numpy
Creación y visualización de datos	Matplotlib
Generación de gráficos	Seaborn
Interacción con widgets para crear la interfaz	Ipywidgets
Visualización de resultados	IPython.display

4. DESARROLLO DEL SISTEMA

En este apartado se explicará cómo se ha diseñado e implementado el sistema desarrollado para el análisis de la concesión de hipotecas. En este punto se detallará cómo funciona la interfaz del usuario, las funcionalidades principales y por último la ejecución del código.

4.1. Interfaz de usuario

Para conseguir una mayor sencillez para el usuario y una mayor facilidad visual, se ha decidido crear una interfaz usando la biblioteca “*Ipywidgets*”. De esta forma, se pueden crear controles dinámicos que facilitan al usuario entender los resultados, sin necesidad de utilizar programación avanzada o alguna herramienta externa.

La interfaz dispone de los siguientes elementos interactivos:

- Desplegable de selección del perfil

Para poder hacer la comparación sobre cuál es la comunidad autónoma con una mayor probabilidad de recibir una hipoteca, primero hay que saber el perfil de la persona que va a solicitarla. Para simplificar este trabajo, el perfil de cada comunidad está ya definido y es básicamente un perfil medio de cada comunidad. Para elaborarlo se ha extraído la información de la edad media de cada comunidad autónoma del INE ya que está actualizado hasta 2024, mientras que los sueldos medios se han seleccionado de ChatGPT debido a que el INE solo está actualizado hasta 2022 mientras que ChatGPT hasta 2024.

Para poder seleccionar este perfil, se ha creado un desplegable en el que se puede elegir una comunidad autónoma que representa el perfil medio de esa comunidad.

A continuación, se muestra la tabla con la información relativa a los perfiles medios que se han explicado anteriormente.

Tabla 2: Información relativa a los perfiles medios

CCAA	SUELDO MEDIO	EDAD MEDIA
Andalucía	21.200	43
Aragón	25.800	45
Asturias	24.500	49
Baleares	22.800	42
Canarias	21.300	44
Cantabria	24.200	46
Castilla y León	23.700	48
Castilla-La Mancha	22.400	44
Cataluña	28.000	43
Valencia	24.500	44
Extremadura	21.000	46
Galicia	23.500	48
Madrid	32.000	43
Murcia	22.300	42
Navarra	28.500	44
País Vasco	30.000	46
La Rioja	24.000	45
Ceuta y Melilla	21.200	38

- Desplegables para la elección de las comunidades autónomas

Además de saber el perfil del usuario que va a solicitar la hipoteca, también se necesita seleccionar qué dos comunidades se quieren estudiar. Por este motivo, se han creado dos desplegables que contienen los nombres de todas las comunidades para que así el usuario pueda elegir y le resulte lo más sencillo posible.

- Botón para calcular en qué comunidad autónoma es más probable la concesión de la hipoteca

Una vez el usuario ha definido el perfil y las comunidades que quiere comparar solamente clicando en el botón de “Calcular”, se ejecuta el código e imprime por pantalla el resultado correspondiente.

- Botón para mostrar los gráficos

Como se ha explicado en puntos anteriores, el trabajo además de calcular en qué comunidad es más probable recibir una hipoteca según su perfil, también contiene un estudio con distintas gráficas de dispersión y una matriz de correlación sobre la concesión de hipotecas históricamente en las distintas comunidades autónomas. Para pedirle al sistema que imprima por pantalla las distintas gráficas, se ha creado el botón de “Mostrar Gráficos”.

Cabe recalcar que ambos estudios son completamente independientes y para este caso no hace falta seleccionar ningún perfil ni comunidad autónoma para poder visualizar las gráficas. Por este motivo se han creado dos botones diferentes y así el usuario puede elegir qué opción desea.

- Botón para mostrar matriz de scores

Además, de los botones recién mencionados se ha incluido en una segunda celda de código la opción de poder mostrar una matriz de scores que compare todos los perfiles con todas las comunidades autónomas y así poder hacer una comparativa más completa.

4.2. Funcionalidades principales

Hasta ahora solamente se ha explicado la interfaz de usuario por lo que en este apartado se explicará más en detalle cómo funciona el código.

4.2.1. Comparación de comunidades autónomas

La comparación de comunidades autónomas se realiza a través de la función

“def calcular_probabilidad(perfil, comunidad1, comunidad2):”.

El código sigue un procedimiento estructurado para poder llevar a cabo este análisis.

El primer paso consiste en la recopilación de datos del usuario, el programa selecciona el sueldo medio y la edad del perfil seleccionado y guarda las comunidades autónomas también seleccionadas. Si por algún casual hay algún fallo y no se ha seleccionado alguna de las dos comunidades, el programa mostrará por pantalla un mensaje para recordarle al usuario que tiene que seleccionar ambas comunidades.

El segundo paso consiste en extraer los datos de la base de datos para poder realizar el análisis con la información proporcionada por el usuario, para ello se utiliza la siguiente fórmula:

```
“datos_com1 = df[(df['CCAA'] == comunidad1) & (df['Año'] >= 2014) & (df['Año'] <= 2023)]”
```

```
“datos_com2 = df[(df['CCAA'] == comunidad2) & (df['Año'] >= 2014) & (df['Año'] <= 2023)]”.
```

Si hubiese un imprevisto y los datos de alguna comunidad no estuvieran disponibles, el programa avisará al usuario con un mensaje.

Una vez se han extraído los datos de la base de datos, queda calcular las variables relevantes, en este caso serán el sueldo, la tasa de desempleo, el número de hipotecas y la edad.

Para el sueldo se utilizarán las siguientes fórmulas:

```
“sueldo_com1 = datos_com1['Sueldo medio'].mean() sueldo_com2 = datos_com2['Sueldo medio'].mean()”,
```

que hace la media de los sueldos de las comunidades autónomas seleccionadas.

Para la tasa de desempleo se utilizarán las siguientes fórmulas:

```
“tasa_com1 = datos_com1['Tasa de desempleo de adultos normalizado'].mean() if edad >= 26 else datos_com1['Tasa de desempleo de jovenes normalizado'].mean()”
```

```
“tasa_com2 = datos_com2['Tasa de desempleo de adultos normalizado'].mean() if edad >= 26 else datos_com2['Tasa de desempleo de jovenes normalizado'].mean()”,
```

que hace exactamente lo mismo que el sueldo, calcula la media de los diferentes años de las comunidades seleccionadas teniendo en cuenta la edad del usuario.

En cuanto a las hipotecas per cápita, se realiza lo mismo, la media del número de hipotecas per cápita de las comunidades.

Por último, para la variable de edad, se ha creado el siguiente bucle: asignando un valor de 1 a las personas menores de 35 años, un valor de 1,5 a las personas entre 35 y 45 años, y un valor de 0,5 a las personas mayores de 45 años.

Ilustración 1: Cálculo de Sedad

```
# Calcular Sedad
if edad < 35:
    sedad = 1
elif 35 <= edad <= 45:
    sedad = 1.5
else:
    sedad = 0.5
```

Habiendo ya calculado las variables relevantes, solo quedaría calcular el “Score” de cada comunidad, cuyas fórmulas son:

Ilustración 2: Cálculo de Score

```
score_com1 = (0.55 * (sueldo / sueldo_com1)) + (0.15 * (1-tasa_com1)) + (0.05 * 100 * hipotecas_com1) + (0.25 * sedad)
score_com2 = (0.55 * (sueldo / sueldo_com2)) + (0.15 * (1-tasa_com2)) + (0.05 * 100 * hipotecas_com2) + (0.25 * sedad)
```

Una vez se ha explicado cómo funciona el código, ahora quedaría explicar las razones teóricas por las que se ha decidido usarla.

La fórmula

$$\text{Score} = (0.55 * (\text{Sueldo personal} / \text{Sueldo medio de la comunidad})) + (0.15 * (1 - \text{Tasa de desempleo normalizada})) + (0.05 * 100 * \text{Hipotecas per cápita normalizado}) + (0.25 * \text{Sedad})$$

está diseñada para comparar la probabilidad de que un solicitante de hipoteca reciba la aprobación. Cada componente tiene un peso asignado basado en su relevancia dentro de los criterios que suelen utilizar las entidades financieras para tomar decisiones sobre préstamos hipotecarios (teniendo en cuenta solamente las variables públicas disponibles). Estos porcentajes han sido propuestos por un profesional con más de 30 años en el sector financiero como una aproximación razonable, basada en su experiencia y en información pública. Es importante señalar que este modelo representa el máximo nivel de detalle que puede extraerse a partir de fuentes disponibles en internet, ya que los bancos no divulgan específicamente cuáles son los factores y ponderaciones exactas que aplican en sus procesos internos de evaluación. A continuación, se analiza y justifica cada término de la fórmula.

- El término

$$“(0.55 * (\text{Sueldo personal} / \text{Sueldo medio de la comunidad}))”$$

tiene el mayor peso dentro de la fórmula, reflejando la importancia crucial de los ingresos personales en la capacidad del solicitante para asumir el compromiso financiero que implica una hipoteca. Tal y como menciona (ING, s.f.), las dos variables clave para comprobar qué hipoteca te puedes permitir son el sueldo y los ahorros, por este motivo el sueldo debería tener el mayor peso de la fórmula. Además, los ingresos altos proporcionan confianza a las entidades financieras, ya que indican una mayor capacidad para afrontar pagos incluso frente a imprevistos, como un aumento en los tipos de interés. La comparación del sueldo personal con el sueldo medio de la comunidad autónoma es especialmente relevante, ya que contextualiza los ingresos del solicitante en función de los estándares locales, lo que permite valorar la estabilidad económica relativa del individuo respecto a su entorno.

- Por su parte, el término

$$“(0.15 * (1 - \text{Tasa de desempleo normalizada}))”$$

aporta un análisis del entorno económico en el que se sitúa el solicitante, considerando que la estabilidad laboral es un factor indirecto pero significativo en la evaluación de la concesión de hipotecas. Una baja tasa de desempleo indica mercados laborales fuertes y reduce el riesgo de que el usuario pierda su trabajo. En cambio, tasas elevadas de desempleo indican una incertidumbre y un mayor riesgo laboral, por lo que podría afectar a la capacidad de pago del usuario al banco. El uso de “*1 - tasa de desempleo normalizada*” transforma este indicador en una métrica positiva dentro de la fórmula, para que así se priorice a las comunidades con mayor tasa de empleo.

Esta tasa de desempleo se ha normalizado para asegurar su comparabilidad con el resto de los términos.

- El componente

$$“(0.05 * 100 * \text{Hipotecas per capita normalizado})”$$

aunque con un peso menor, aporta información valiosa sobre las políticas crediticias locales y la accesibilidad al crédito en la región del solicitante. Esta variable mide la proporción de hipotecas concedidas por habitante, reflejando la flexibilidad de los bancos en esa comunidad. Una alta proporción sugiere políticas de crédito más accesibles y una economía local favorable, mientras que un bajo valor puede ser indicativo de restricciones

crediticias. La normalización del dato asegura su comparabilidad con los demás términos de la fórmula, mientras que la multiplicación por 100 amplifica su impacto para mantenerlo en una escala proporcional al resto de los factores evaluados.

- Finalmente, el término:

$$“(0.25 * \text{Edad})”$$

introduce la variable de la edad del solicitante, que, aunque no es el factor más determinante, influye significativamente en la duración y el riesgo del crédito. Las entidades financieras prefieren conceder hipotecas a personas más jóvenes, ya que tienen más tiempo para devolver el préstamo antes de alcanzar la jubilación, lo que permite establecer plazos más largos y reduce el riesgo de impago. Además, las personas jóvenes tienden a aumentar sus ingresos a lo largo de su carrera profesional, lo que refuerza su capacidad de pago. La fórmula categoriza la edad en tres rangos, asignando un valor de 1 a los menores de 35 años, quienes disfrutan de mayor flexibilidad para obtener plazos largos; un valor de 1.5 a las personas entre 35 y 45 años, que representan un perfil de estabilidad laboral y solvencia consolidada; y un valor de 0.5 a los mayores de 45 años, quienes enfrentan mayores restricciones debido a la cercanía de la jubilación y la limitación en los plazos de devolución. (Selectra, 2025)

En resumen, la fórmula recién explicada combina de manera coherente factores individuales y contextuales que influyen en la evaluación del riesgo crediticio. La ponderación de cada componente refleja su relevancia dentro de las prácticas estándar de las entidades financieras, proporcionando una herramienta equilibrada y razonable para poder calcular en qué comunidad es más fácil para el usuario recibir la hipoteca.

4.2.2. Generación de gráficos

El objetivo principal de la generación de los gráficos dinámicos es poder analizar visualmente cómo se comporta cada comunidad autónoma ante las variables previamente mencionadas.

Para poder generar estos gráficos se han utilizado las bibliotecas llamadas “*Matplotlib*” y “*Seaborn*”, ya que permiten visualizaciones claras que facilitan el entendimiento de las gráficas.

Hay una función específica que se encarga de este punto llamada:

“def mostrar_graficos()”,

la cual empieza a funcionar cuando el usuario pulsa el botón de *“Mostrar gráficos”* de la interfaz. En este caso, no hace falta añadir ningún dato manualmente, ya que el programa examina todas las comunidades autónomas por igual.

El flujo del código es el siguiente:

- Primero se definen las variables que se consideran de interés, en este caso las variables son *“PIB per cápita”*, *“sueldo medio”* y *“tasa de desempleo”*. Estas tres variables son las que se utilizarán en el eje x y las que se enfrentarán a los valores de *“hipotecas per cápita”* e *“importe de hipotecas per cápita”*, que irán en el eje y de las gráficas.

- Seguidamente se generarán dos diagramas de dispersión por cada una de las tres variables del eje x (una enfrentándose a *“hipotecas per cápita”* y otra enfrentándose a *“importe de hipotecas per cápita”*).

En primer lugar, la función *“plt.figure(figsize=(8, 5))”* marca el tamaño que tendrá la gráfica en la interfaz (8 unidades de ancho y 5 unidades de alto). Luego el código crea la gráfica de dispersión con *“sns.scatterplot()”*, donde en el eje x se colocan las variables explicadas en el párrafo anterior y en el eje y se colocan las variables hipotecas per cápita o importe de hipotecas per cápita, para poder diferenciar las comunidades autónomas, el parámetro *“hue=df[‘CCAA’]”* usando la paleta *“Set2”* pone cada comunidad de un color diferente. A continuación, se le añade un título a cada gráfica y a los ejes con los parámetros, *“plt.title()”*, *“plt.xlabel()”* y *“plt.ylabel()”*. Además, se coloca una leyenda utilizando *“plt.legend(bbox_to_anchor=(1.05, 1), loc=‘upper left’)”* y por último, se utiliza para mostrar el gráfico por pantalla el parámetro *“plt.show()”*.

- Una vez se han generado las 6 gráficas de dispersión, solo quedaría mostrar la matriz de correlaciones con su mapa de calor.

Para poder crearlo, primero se seleccionan las columnas de las variables que queremos tener en la matriz, en este caso son las 5 variables que aparecen en las gráficas de dispersión junto con el parámetro *.corr()* para que de esta manera el código devuelva los coeficientes de Pearson, explicados anteriormente, para cada una de las variables. A continuación, se ejecuta la función

```
“sns.heatmap(correlaciones, annot=True, cmap='coolwarm', fmt=".2f")”
```

para crear el mapa de calor, gracias a

```
“annot=True y fmt=".2f”
```

los valores numéricos de las correlaciones se pueden observar dentro de sus correspondientes celdas y *“cmap='coolwarm'”* resalta las correlaciones positivas en rojo y las negativas en azul para que así sea mucho más sencillo. Por último, solo quedaría ponerle un título a la gráfica y mostrarla por pantalla con la función *“plt.show()”*.

4.2.3. Matriz de scores

Además de las funcionalidades previamente explicadas, se ha añadido la posibilidad de que el usuario pueda comparar todos los perfiles con todas las comunidades autónomas y así poder tener una visión mas genérica.

El funcionamiento de esta matriz es idéntico al apartado “Comparación de comunidades autónomas”, con la única diferencia de que ahora se enfrentan todas las comunidades con todos los perfiles.

4.3. Requisitos previos para la ejecución del código

Antes de ejecutar el código hay que tener en cuenta ciertos requisitos previos para que así funcione correctamente y no ocurra ningún fallo. Los requisitos son los siguientes:

- Acceso a Google Colab, se puede utilizar desde el navegador.
- Tener cargada la base de datos en el mismo entorno de Google Colab que el código.

5. RESULTADOS

En este apartado se mostrarán y comentarán los resultados obtenidos a partir del análisis de los datos y de los modelos estadísticos y visuales. Se presentarán tanto las gráficas y sus explicaciones como las principales conclusiones obtenidas del modelo que compara las comunidades autónomas según el perfil.

5.1. Comparación de comunidades autónomas

En este apartado se comentarán los resultados obtenidos con el modo de calcular en qué comunidad es más probable recibir una hipoteca.

Lo primero que ve el usuario son los 3 desplegados y los 2 botones:

Ilustración 3: Desplegables y botones de la interfaz

Selecciona un perfil y dos comunidades autónomas para la comparación:

Perfil:	Andalucía	▼
Comunidad ...	Andalucia	▼
Comunidad ...	Andalucia	▼
<button>Mostrar Gráficos</button>		
<button>Calcular</button>		

Una vez selecciona el perfil que quiere estudiar, le salta un mensaje indicándole el perfil que ha seleccionado, por ejemplo, en este caso se ha elegido Cataluña.

Ilustración 4: Ejemplo de perfil seleccionado

Perfil seleccionado: Edad media: 43, Sueldo bruto anual: 28000

Ahora imaginemos que el usuario quiere comparar donde sería para él más fácil obtener una hipoteca si en su ciudad, Cataluña, o en Andalucía, por lo que selecciona esas dos comunidades y clicla en “Calcular”, y el resultado que obtiene es lo siguiente:

Ilustración 5: Ejemplo de comunidades y perfil a analizar

Selecciona un perfil y dos comunidades autónomas para la comparación:

Perfil:	Cataluña	▼
Comunidad ...	Cataluña	▼
Comunidad ...	Andalucia	▼
Mostrar Gráficos		
Calcular		

Con el perfil seleccionado, es más probable obtener una hipoteca en Andalucía.
Score Cataluña: 1.80, Score Andalucía: 2.23.

Lo cual tiene sentido ya que el 55% del peso de la fórmula tiene que ver con el sueldo, y en este caso el sueldo del perfil es bastante más alto que el sueldo medio de Andalucía. Es cierto que Andalucía tiene mayor tasa de desempleo que Cataluña por lo que podría ser más difícil poder acceder a una vivienda, pero el peso de esta variable es solamente de un 15%.

Hasta ahora simplemente hemos visto como un perfil de Cataluña influye en su propia comunidad autónoma y en Andalucía, ahora vamos a comparar dos perfiles diferentes en estas dos mismas comunidades autónomas como son Cataluña y Andalucía.

Para el siguiente estudio se han seleccionado las comunidades de La Rioja y Castilla y León ya que tienen sueldos parecidos, el perfil de La Rioja tiene una edad media 44 años y una sueldo de 24.000 euros brutos anuales, mientras que el perfil de Castilla y León tiene una edad media de 46 años y un sueldo de 23.700 euros brutos anuales.

Estos son los resultados obtenidos:

Ilustración 6: Ejemplo de comunidades y perfil a analizar

Selecciona un perfil y dos comunidades autónomas para la comparación:

Perfil:	La Rioja	▼
Comunidad ...	Cataluña	▼
Comunidad ...	Andalucía	▼
Mostrar Gráficos Calcular		

Con el perfil seleccionado, es más probable obtener una hipoteca en Andalucía. Score Cataluña: 1.71, Score Andalucía: 2.12.

Ilustración 7: Ejemplo de comunidades y perfil a analizar

Selecciona un perfil y dos comunidades autónomas para la comparación:

Perfil:	Castilla y León	▼
Comunidad ...	Cataluña	▼
Comunidad ...	Andalucía	▼
Mostrar Gráficos Calcular		

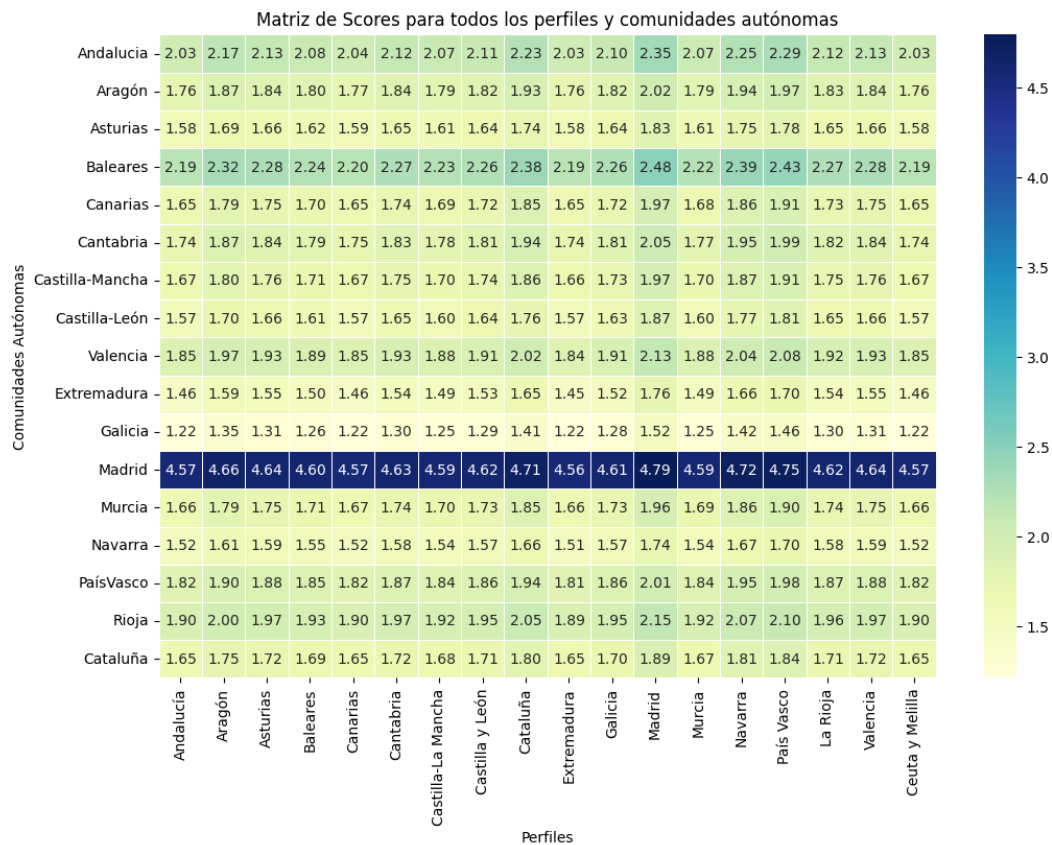
Con el perfil seleccionado, es más probable obtener una hipoteca en Andalucía. Score Cataluña: 1.46, Score Andalucía: 1.86.

Con estos resultados podemos observar como las probabilidades de obtener una hipoteca siendo de La Rioja son más altas que de Castilla León, principalmente por el hecho de que la edad media de La Rioja es menor de 45 años y la de Castilla y León mayor. Este motivo hace que el 25% de peso que tiene la edad incline la balanza a favor de La Rioja, ya que como se ha comentado en apartados anteriores, a partir de 45 años los bancos establecen mayores dificultades para pedir una hipoteca.

En conclusión, como se ha podido ver con los ejemplos seleccionados, este modo permite comparar y decir qué comunidad es más probable de recibir una hipoteca en función del perfil que se seleccione.

Para poder entender mejor cómo funcionan los scores y poder comparar todos los perfiles con las diferentes comunidades autónomas se ha creado la matriz adjuntada a continuación.

Ilustración 8: Matriz de scores para todas las CCAA



En esta matriz, el eje de las columnas representa los perfiles de las comunidades autónomas y el eje de las filas representa a cada comunidad autónoma. Es decir, para saber en qué comunidad es más probable recibir una hipoteca para cierto perfil, sería suficiente con encontrar el número más alto en la columna que contenga a ese perfil.

Lo primero que llama la atención según se observa esta gráfica es cómo todos los perfiles son más propensos a recibir una hipoteca en Madrid que en cualquier otra comunidad. Esto se debe a que la tasa de desempleo es inferior y sobre todo debido a que, como se observará en el siguiente punto, tanto el número de hipotecas como el importe de estas hipotecas en Madrid están muy por encima de cualquier otra comunidad. Además, se puede observar cómo dentro de la fila de Madrid, los perfiles con más probabilidad son los propios madrileños, vascos y navarros, que son los que tienen el sueldo más alto.

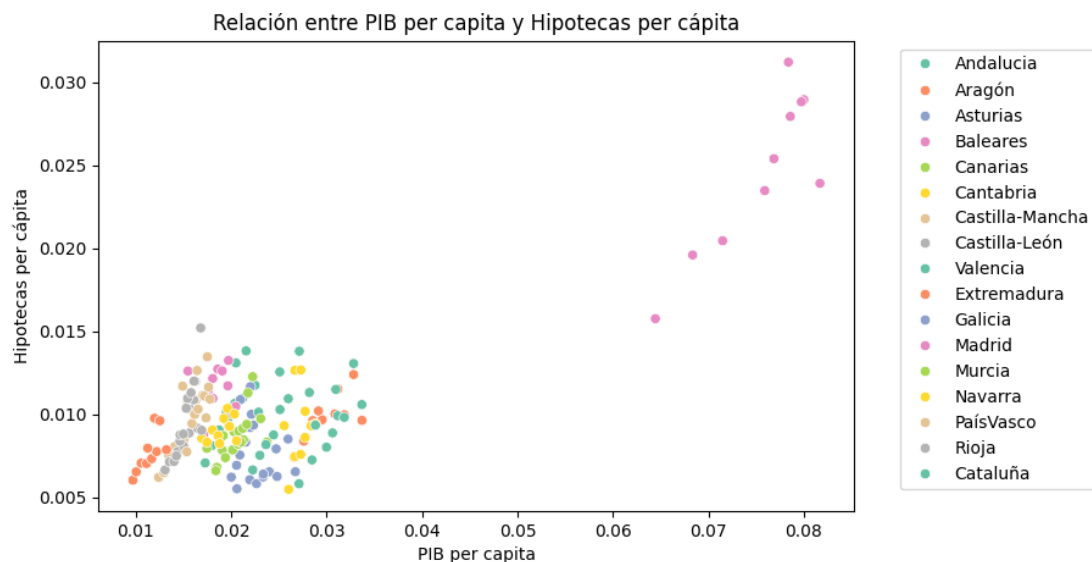
Tal y como se ha estudiado Madrid, se podrían estudiar cada comunidad o cada perfil por separado y así poder llegar a extraer las conclusiones necesarias gracias a este mapa de calor.

5.2. Análisis de las gráficas

Como se ha comentado al inicio de este apartado, en este punto se analizarán las gráficas obtenidas.

5.2.1. Gráficos de Dispersión

Ilustración 9: Relación entre PIB per cápita e Hipotecas per cápita



En esta gráfica podemos observar cómo claramente todas las comunidades tienen valores que se podrían considerar similares, mientras que Madrid está muy por encima del resto en cuanto a hipotecas per cápita y en cuanto a PIB per cápita.

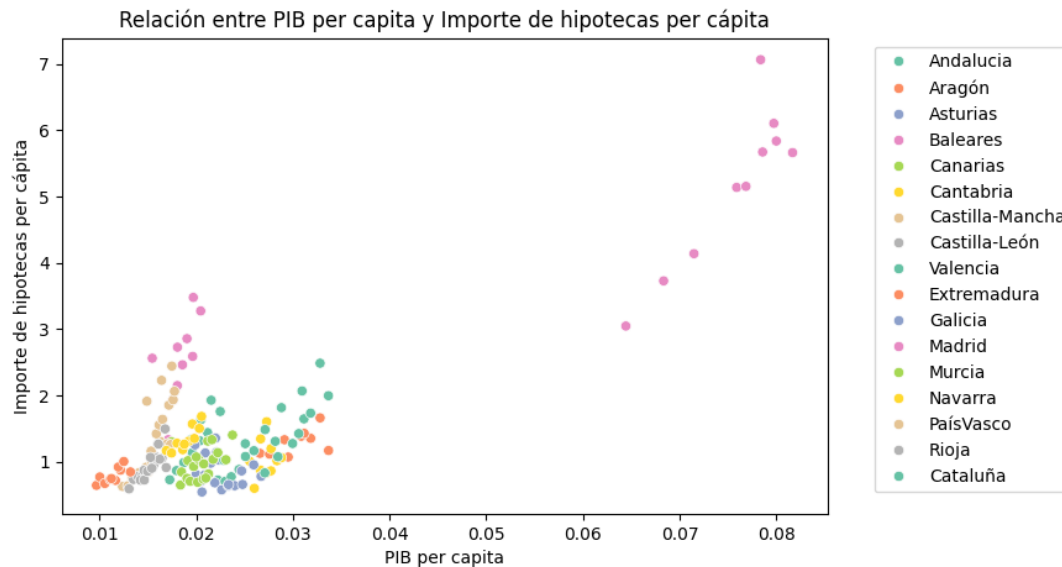
Aparte de Madrid, también se pueden observar datos interesantes como por ejemplo que las comunidades autónomas con mayor PIB per cápita son Cataluña, Aragón y Cantabria y las que menos tienen son Extremadura, Castilla La Mancha y La Rioja.

Otro punto importante para destacar en esta gráfica es que los puntos de cada comunidad autónoma siguen una tendencia de una línea de 45 grados, lo que indica que, a lo largo de los años, crecen de manera parecida (con la escala de los ejes elegida para la gráfica) el PIB per cápita y el número de hipotecas per cápita.

Para finalizar con esta gráfica, con los comentarios que se acaban de argumentar, se puede concluir que, exceptuando Madrid, la relación del PIB per cápita con el número de hipotecas per cápita es significativo, pero solamente dentro de las comunidades. Es decir, cada comunidad funciona de manera independiente y no por tener mayor PIB per cápita que otra comunidad va a recibir más hipotecas, lo que sí que influye es si el PIB per cápita

de la comunidad aumenta, el número de hipotecas que se conceden aumentan también, pero solo comparándolo con su propia comunidad autónoma.

Ilustración 10: Relación entre PIB per cápita e Importe de hipotecas per cápita



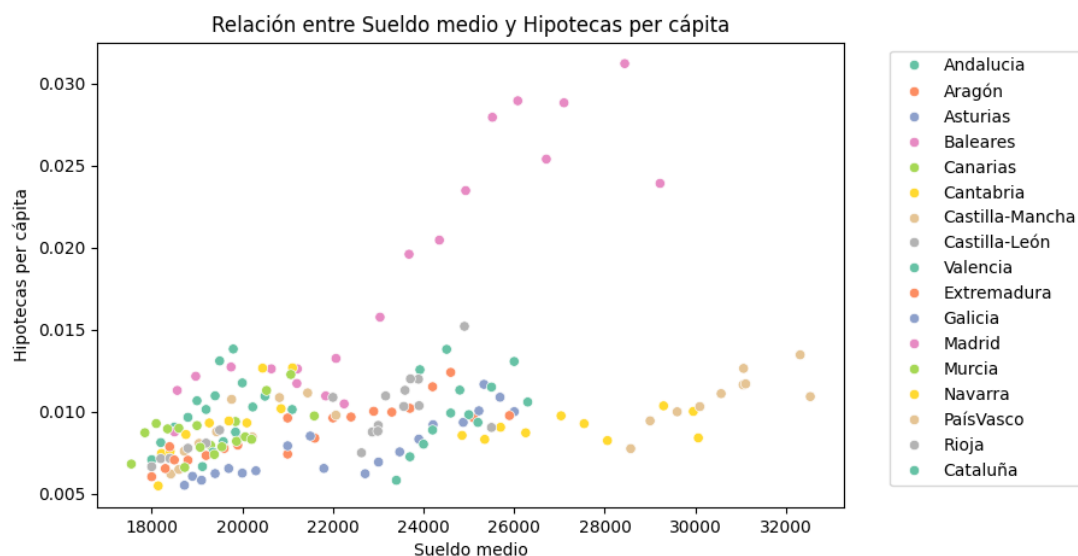
Esta segunda gráfica es parecida a la anterior, con el único matiz de que en este caso se relaciona el PIB per cápita con el importe de hipotecas, no con el número.

Al igual que en la anterior, Madrid está muy por encima del resto en ambos parámetros y el resto de las comunidades están más parejas.

La única diferencia es que en este caso los valores de cada comunidad están más juntos entre sí, lo que indica que, pese a que sigue habiendo una ligera tendencia a aumentar el importe de las hipotecas si el PIB aumenta, no es un cambio tan grande como en la anterior gráfica.

Además, entre comunidades, no hay relación entre el importe de las hipotecas per cápita y el PIB per cápita. Por ejemplo, Baleares tiene menor PIB per cápita que Aragón, pero recibe un mayor importe de hipotecas.

Ilustración 11: Relación entre Sueldo Medio e Hipotecas per cápita

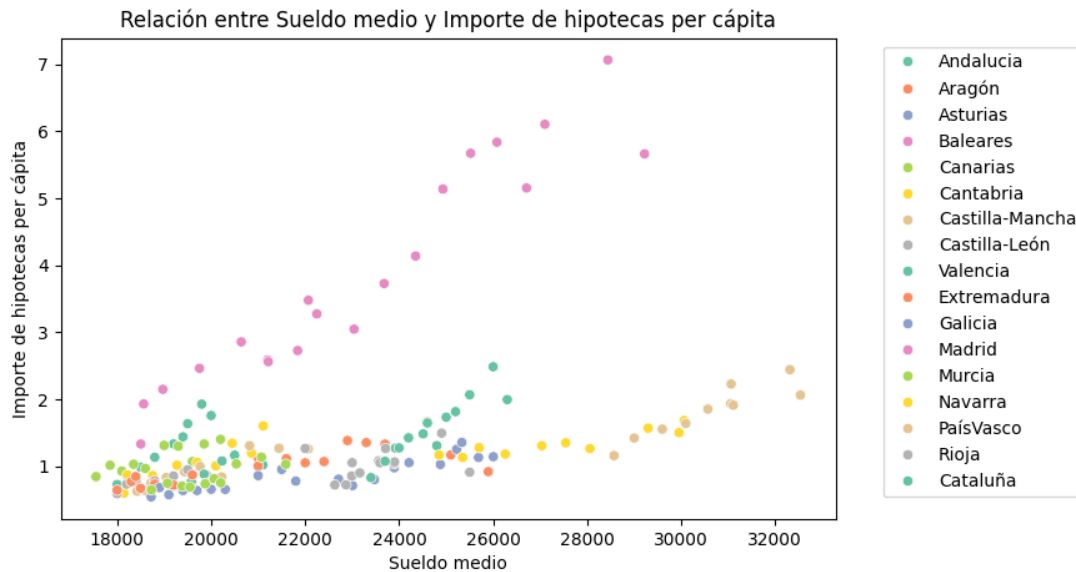


En esta gráfica se relaciona el sueldo medio, con el número de hipotecas per cápita y se puede observar cómo hay una diferencia abismal entre los sueldos de País Vasco y Navarra con los de otras comunidades como Canarias o Extremadura.

Esta diferencia de sueldos no es significativa ya que el número de hipotecas no aumenta si el sueldo medio aumenta. Por ejemplo, Cataluña, pese a que su sueldo medio es inferior al del País Vasco, le conceden un número parecido o incluso mayor de hipotecas.

Otro comentario para tener en cuenta es que, con esta gráfica es posible ver cómo los sueldos medios de las comunidades autónomas avanzan de diferente manera. Por ejemplo, los puntos de Navarra están mucho más “estirados” que los de Murcia, lo que indica que, pese a que tienen mayor sueldo medio, esa diferencia aumenta cada año.

Ilustración 12: Relación entre Sueldo Medio e Importe de hipotecas per cápita

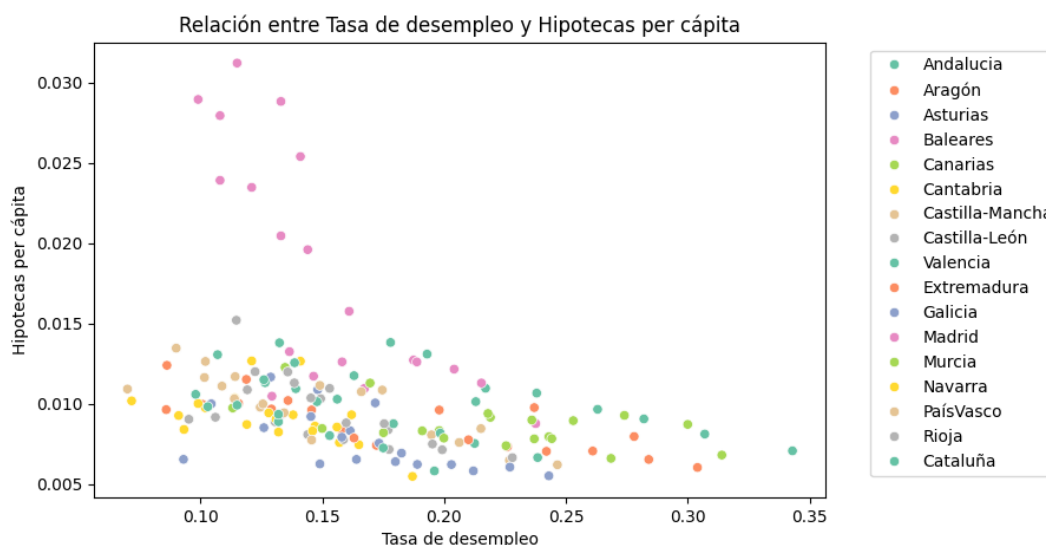


Esta 4ª gráfica es igual que la anterior con la diferencia de que en el eje y estamos teniendo en cuenta el importe de hipotecas per cápita, no el número de hipotecas.

Al igual que la anterior, se puede observar como el sueldo medio no influye en el importe de hipotecas, ya que hay comunidades con sueldos más altos que reciben menos dinero por las hipotecas.

Una comunidad interesante para comentar en esta gráfica es Baleares, en la anterior gráfica se podía observar cómo su sueldo medio y el número de hipotecas per cápita era parecido a comunidades como Andalucía o Valencia, pero en esta gráfica el importe es significativamente mayor, lo que nos indica que las hipotecas que se conceden en Baleares son de un importe mayor a lo que lo hacen en Valencia o Andalucía.

Ilustración 13: Relación entre Tasa de desempleo e Hipotecas per cápita

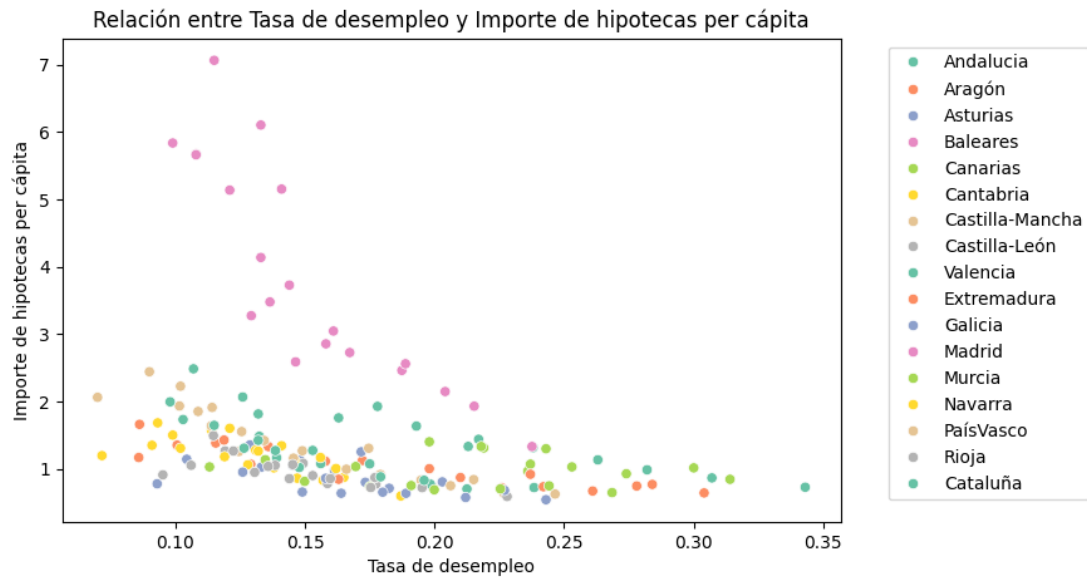


En esta 5ª gráfica se va a relacionar la tasa de desempleo de cada comunidad con el número de hipotecas per cápita. Se puede observar cómo de manera distinta a las anteriores gráficas, esta variable sí que puede estar relacionada con el número de hipotecas que se conceden.

Si exceptuamos Madrid, se puede observar una ligera tendencia de los puntos a disminuir en el eje y en cuanto se aumenta el eje x. Esto indica que cuanto mayor es el desempleo en esa comunidad, más reacios son los bancos a la hora de conceder hipotecas.

Como dato, cabe resaltar que las comunidades con mayor tasa de desempleo son Andalucía, Canarias y Extremadura, mientras que las comunidades con menor tasa de desempleo son País Vasco, Navarra y Aragón. Además, estos valores tienden a disminuir en el eje x a lo largo de los años, lo que significa que en 2014 la tasa de desempleo general era mucho mayor a la de 2023.

Ilustración 14: Relación entre Tasa de desempleo e Importe de hipotecas per cápita

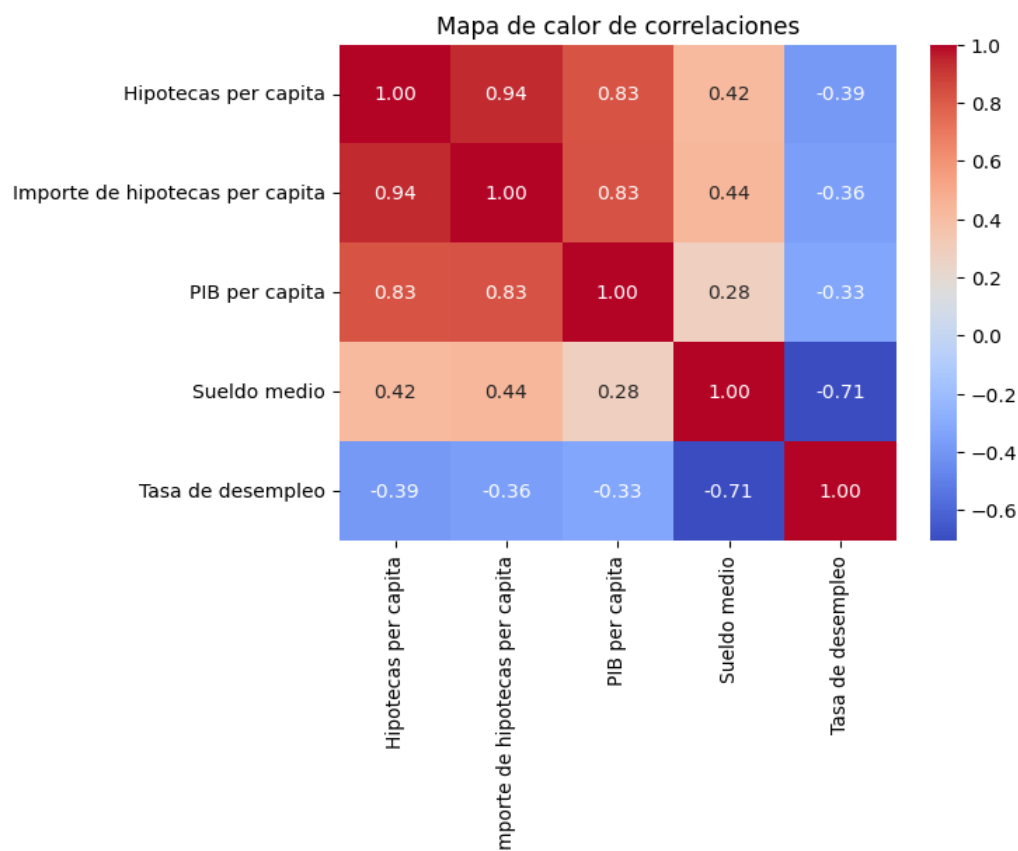


En esta última gráfica de dispersión, al igual que con los anteriores parámetros, lo que se diferencia con la 5ª es que esta vez estamos comparando la tasa de desempleo con el importe de las hipotecas per cápita.

En este caso, quitando Madrid y Baleares, los valores en el eje y son muy parecidos, independientemente del valor en el eje x que tengan, lo que indica que no es significativa la tasa de desempleo en este caso.

5.2.2. Mapa de calor de correlaciones

Ilustración 15: Mapa de calor de correlaciones



Para poder comprender este mapa de calor de correlaciones en relación con la temática abordada en este trabajo hay que fijarse en las dos primeras filas o las dos primeras columnas, ya que la matriz es simétrica. Esto se debe a que en las dos primeras filas o columnas es donde se encuentran las variables de hipotecas per cápita e importe de hipotecas per cápita comparadas con el resto de las variables.

Lo primero que llama la atención es que la relación entre hipotecas per cápita e importe de hipotecas per cápita es cercano a 1 pero sin llegar, lo que hace que tenga sentido las gráficas en las que se resaltaba Baleares que acabamos de observar en el punto anterior.

Otra variable que se puede observar con un alto indicador es el PIB per cápita con respecto a ambas variables de hipotecas con un 83%, lo que nos indica una correlación positiva fuerte.

La siguiente variable es el sueldo medio, que tal y como se ha comentado en las respectivas gráficas, no influye mucho en relación con las variables de las hipotecas, solamente un 42%.

Por último, la variable de la tasa de desempleo, como es lógico, influye de manera negativa ya que cuanta mayor tasa de desempleo hay, menor número de hipotecas se conceden. En este caso, es la variable que menos influye a la hora de conceder hipotecas con un 36% y 39% con respecto al número de hipotecas y al importe de las hipotecas respectivamente.

Otros valores que destacar en esta gráfica que nos sirven como comprobación de que la matriz está bien hecha es la relación entre la tasa de desempleo y el sueldo medio con un -71%, que es un valor relativamente alto, lo que tiene sentido. Y también comentar la relación entre sueldo medio y PIB per cápita que es la menor de toda la matriz solamente con un 28%, que podría ocurrir porque algunas comunidades tienen un PIB alto pero los salarios son más bajos.

6. CONCLUSIONES

En este apartado se comentarán las conclusiones que se han extraído de este trabajo, estas reflejan un análisis sobre la relación de ciertas variables sociales, económicas y financieras con la concesión de hipotecas en España, y más concretamente en las diferentes comunidades autónomas. No solo se comentarán las conclusiones de los resultados obtenidos, sino que también se comentará el marco teórico, la metodología y las implicaciones de los hallazgos.

Durante todo el proyecto se ha expuesto que obtener o no una hipoteca no depende solamente de factores o variables individuales, sino que también depende de temas macroeconómicos, políticas bancarias, etc, y estas variables no individuales no tienen por qué ser iguales en todas las comunidades autónomas.

Desde un punto de vista teórico, este trabajo demuestra que la estabilidad laboral, el riesgo ante el endeudamiento y otros motivos son influyentes en la concesión de hipotecas. También se ha explicado cómo los tipos de interés y evoluciones del mercado como las crisis, son un factor clave para estas concesiones.

En cuanto a la metodología, el haber combinado las fórmulas para calcular en qué comunidad es más probable recibir la hipoteca y las gráficas de dispersión ha permitido identificar patrones y correlaciones entre las variables seleccionadas. Además, el haber obtenido los mismos datos para todas las comunidades ha hecho que los resultados sean más fiables. Sin embargo, es importante reconocer que al no tener acceso a datos cualitativos o a algunos factores que también podrían influir en la concesión de hipotecas, este informe no está completamente exacto.

Sobre los resultados obtenidos, se ha confirmado que tanto el sueldo como la edad de los usuarios son el factor más influyente a la hora de calcular la probabilidad. El análisis de la tasa de desempleo muestra cómo existe una relación negativa con la concesión de hipotecas, y esta relación no es la misma en cada comunidad autónoma, ya que algunas de ellas pese a tener altas tasas de desempleo mostraban altos niveles en las concesiones, lo que indica que otros factores externos también están influyendo.

Otro aspecto clave que se ha identificado es la relación entre las hipotecas y el PIB per cápita, aunque la correlación es positiva como se esperaba, no es determinante cuando se comparan entre comunidades autónomas. Por ejemplo, hay regiones con un PIB per cápita elevado que no muestran un mayor acceso a hipotecas, lo que indica que el sueldo de la

población o la política de los bancos en cada comunidad son factores más influyentes que la propia riqueza de la región.

Además, el uso de una matriz de calor comparando todos los perfiles y comunidades autónomas ha permitido identificar visualmente en qué comunidades es más probable recibir una hipoteca según el perfil. Destaca Madrid como la comunidad con mayor probabilidad para todos los perfiles, debido a su baja tasa de desempleo y al elevado número e importe de hipotecas. También se observa que los perfiles más favorecidos en Madrid son los madrileños, vascos y navarros, lo que confirma la influencia del nivel salarial.

Por otra parte, con este estudio se ha podido observar cómo el mercado hipotecario español está en gran medida marcado por factores europeos del Banco Central Europeo (BCE). En los próximos años, el auge de nuevas formas de acceder a esta financiación puede cambiar el rumbo de este sector.

En conclusión, con este proyecto no solo se puede entender cuánto afectan ciertos factores o variables a la concesión de hipotecas, sino que también se puede utilizar para futuras investigaciones que profundicen en las desigualdades regionales o en las políticas económicas y financieras. Estos resultados obtenidos pueden ser útiles sobretudo para usuarios que quieran ver qué factores influyen más y donde es más probable que reciban una hipoteca con sus datos y también puede servir para entidades bancarias para que garanticen un mercado hipotecario más equitativo para cada comunidad autónoma.

7. LIMITACIONES Y FUTURAS LÍNEAS DE INVESTIGACIÓN

A pesar de haber intentado lograr la máxima exactitud en los cálculos para obtener resultados fiables con la información disponible, es necesario reconocer que esta no es completamente precisa, por lo que debe tenerse en cuenta esta limitación. Además, cabe mencionar que la literatura académica es escasa y el trabajo se ha basado en su mayoría en artículos profesionales.

El principal punto de mejora reside en los datos, este proyecto se ha efectuado contando únicamente con las variables de PIB, habitantes, sueldo, tasa de desempleo, edad, número de hipotecas e importe de hipotecas. Estos son los datos disponibles en internet que se pueden descargar. Evidentemente, hay más variables que tienen en cuenta los bancos a la hora de conceder hipotecas, como por ejemplo el dinero ahorrado o propiedades que posean.

En este trabajo se ha intentado ver si hubiese alguna correlación entre las variables y en qué comunidad es más probable recibir la hipoteca, por lo que se ha supuesto que el dinero ahorrado es el mismo en todas las situaciones.

Como consecuencia de esta ausencia de datos relevantes no se ha conseguido obtener cual es la probabilidad de poder recibir una hipoteca en cada comunidad autónoma, lo que sí que se puede es ver cuál es más probable, pero no la probabilidad exacta.

Otro posible punto de mejora es que los perfiles de cada comunidad autónoma ya están definidos, podría ser más realista si el usuario pudiera meter sus propios datos y no tener que seleccionar uno de los que ya están predefinidos. El motivo por el que se decidió no añadir esta función fue para simplificar la programación ya que se consideró que el usuario no introdujese sus propios valores no aportaba valor a los objetivos de este trabajo.

Un factor muy importante que tampoco se ha considerado es la inflación y los tipos de interés, ya que son factores clave para la concesión de hipotecas ya que en la mayoría de los casos son valores muy importantes para decidir cuál será la cantidad que el usuario deberá pagar al banco por prestarle el dinero necesario.

En cuanto a futuras líneas de investigación, este mismo análisis se podría profundizar a nivel provincial o incluso municipal, hay en comunidades autónomas en las que habrá pocas diferencias, pero otras como Madrid o Cataluña, tienen dinámicas inmobiliarias

muy diferentes a otros municipios de su propia comunidad o provincia. Además, este análisis se podría extrapolar a países, y así poder comparar España con el resto de los países de la Unión Europea o incluso de fuera de Europa, para observar qué diferencias hay entre países.

Como se ha mencionado anteriormente, no se han considerado la inflación ni los tipos de interés, factores clave determinados por el Banco Central Europeo (BCE). En futuras investigaciones, sería interesante analizar cómo estos cambios afectan la concesión de hipotecas en nuestro país.

Por último, una posible línea de investigación que se podría seguir es la utilización de “machine learning” para predecir cómo estos patrones influyen en la probabilidad de recibir una hipoteca.

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

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

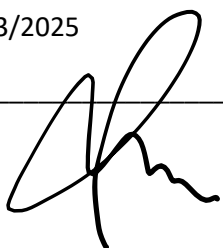
Por la presente, yo, Paula López de Uralde Lozano, estudiante de E2+Business Analytics de la Universidad Pontificia Comillas al presentar mi Trabajo Fin de Grado titulado " Hipotecas y Desigualdad Regional en España: Desarrollo de un Modelo Analítico para Evaluar el Acceso a la Financiación", declaro que he utilizado la herramienta de Inteligencia Artificial Generativa ChatGPT u otras similares de IAG de código sólo en el contexto de las actividades descritas a continuación [el alumno debe mantener solo aquellas en las que se ha usado ChatGPT o similares y borrar el resto. Si no se ha usado ninguna, borrar todas y escribir “no he usado ninguna”]:

1. **Brainstorming de ideas de investigación:** Utilizado para idear y esbozar posibles áreas de investigación.
2. **Referencias:** Usado juntamente con otras herramientas, como Science, para identificar referencias preliminares que luego he contrastado y validado.
3. **Metodólogo:** Para descubrir métodos aplicables a problemas específicos de investigación.
4. **Interpretador de código:** Para realizar análisis de datos preliminares.
5. **Estudios multidisciplinares:** Para comprender perspectivas de otras comunidades sobre temas de naturaleza multidisciplinar.
6. **Constructor de plantillas:** Para diseñar formatos específicos para secciones del trabajo.
7. **Corrector de estilo literario y de lenguaje:** Para mejorar la calidad lingüística y estilística del texto.
8. **Generador de datos sintéticos de prueba:** Para la creación de conjuntos de datos ficticios.
9. **Generador de problemas de ejemplo:** Para ilustrar conceptos y técnicas.
10. **Revisor:** Para recibir sugerencias sobre cómo mejorar y perfeccionar el trabajo con diferentes niveles de exigencia.
11. **Traductor:** Para traducir textos de un lenguaje a otro.

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

Fecha: 22/03/2025

Firma: _____



Bibliografía

- (UCI), U. d. (2025). *10 tendencias del sector inmobiliario en 2025*.
- 20minutos. (6 de Octubre de 2024). Mapa de la desigualdad en vivienda: la hipoteca media en Baleares triplica la de Lugo o Ciudad Real.
- Balderix, A. (s.f.). *Probabilidad y Estadística*. Obtenido de <https://www.probabilidadyestadistica.net/diagrama-de-dispersion/>
- BBVA. (2021). *Evolución reciente del mercado hipotecario*.
- BBVA. (s.f.). *BBVA*. Obtenido de <https://www.bbva.es/finanzas-vistazo/ef/hipotecas/como-funciona-hipoteca.html>
- ING. (s.f.). *ING*. Obtenido de ING: <https://www.ing.es/enararanja/comprar-casa/que-hipoteca-me-puedo-permitir/>
- LearnStatistics. (s.f.). *Learn Statistics Easily*. Obtenido de <https://es.statisticseasily.com/glosario/%C2%BFQu%C3%A9-es-la-matriz-de-correlaci%C3%B3n-explicada-en-detalle%3F/>
- MontePrincipe Gestión. (2025). Obtenido de <https://www.montepincipegestion.es/ley-de-contratos-de-credito-inmobiliario-o-lcci/>
- Press, E. (2 de Enero de 2025). La concesión de hipotecas podría crecer un 10% en 2025, hasta alcanzar las 400.000 operaciones, según UCI. *Europa Press*.
- Research, B. (Febrero de 2021). *Evolución reciente del mercado hipotecario*.
- Selectra. (17 de Enero de 2025). *Selectra*. Obtenido de <https://selectra.es/finanzas/hipotecas/edad-maxima-hipoteca>
- SER, R. C. (10 de Enero de 2025). *Cadena SER*. Obtenido de <https://cadenaser.com/nacional/2025/01/10/la-vivienda-se-encarece-en-espana-mas-del-doble-que-en-la-union-europea-cadena-ser/#:~:text=La%20vivienda%20se%20encareci%C3%B3%20en%20Espa%C3%B1a%20un%208%2C4%25,precios%20en%20el%20conjunto%20de%20la%20Uni%C3%B3n>
- Soza Marañón, A. (2019). *ANÁLISIS ECONOMETRICO SOBRE LA DEMANDA DE HIPOTECAS*. Trabajo Fin de Grado.
- Vieira, D. (15 de Abril de 2024). *HostGator*. Obtenido de <https://www.hostgator.mx/blog/python-como-usa-lenguaje-programacion/>
- Vieira, D. (5 de Abril de 2024). *HostGator*. Obtenido de <https://www.hostgator.mx/blog/google-colab/>

Anexo I: Código de programación

```
# Importar bibliotecas
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import ipywidgets as widgets
from IPython.display import display, clear_output

# Cargar la base de datos desde el archivo de Excel
df = pd.read_excel('BBDD_TFG.xlsx')

# Crear las columnas necesarias si no están creadas
if 'Hipotecas per capita' not in df.columns:
    df['Hipotecas per capita'] = df['Número de hipotecas'] /
df['Población']

if 'Importe de hipotecas per capita' not in df.columns:
    df['Importe de hipotecas per capita'] = df['Importe de
hipotecas'] / df['Población']

# Normalizar columnas
def normalizar_columna(columna):
    return (columna - columna.min()) / (columna.max() -
columna.min())

df['Hipotecas per capita normalizado'] =
normalizar_columna(df['Hipotecas per capita'])
df['Sueldo medio normalizado'] = normalizar_columna(df['Sueldo
medio'])
df['Tasa de desempleo de jovenes normalizado'] = 1 -
normalizar_columna(df['Tasa de desempleo de jovenes'])
df['Tasa de desempleo de adultos normalizado'] = 1 -
normalizar_columna(df['Tasa de desempleo de adultos'])

# Perfiles medios por CCAA
perfiles = {
    "Andalucía": {"Edad media": 43, "Sueldo bruto anual": 21200},
    "Aragón": {"Edad media": 45, "Sueldo bruto anual": 25800},
    "Asturias": {"Edad media": 49, "Sueldo bruto anual": 24500},
    "Balears": {"Edad media": 42, "Sueldo bruto anual": 22800},
    "Canarias": {"Edad media": 44, "Sueldo bruto anual": 21300},
    "Cantabria": {"Edad media": 46, "Sueldo bruto anual": 24200},
    "Castilla-La Mancha": {"Edad media": 44, "Sueldo bruto anual":
22400},
    "Castilla y León": {"Edad media": 48, "Sueldo bruto anual":
23700},
```

```

    "Cataluña": {"Edad media": 43, "Sueldo bruto anual": 28000},
    "Extremadura": {"Edad media": 46, "Sueldo bruto anual": 21000},
    "Galicia": {"Edad media": 48, "Sueldo bruto anual": 23500},
    "Madrid": {"Edad media": 43, "Sueldo bruto anual": 32000},
    "Murcia": {"Edad media": 42, "Sueldo bruto anual": 22300},
    "Navarra": {"Edad media": 44, "Sueldo bruto anual": 28500},
    "País Vasco": {"Edad media": 46, "Sueldo bruto anual": 30000},
    "La Rioja": {"Edad media": 45, "Sueldo bruto anual": 24000},
    "Valencia": {"Edad media": 44, "Sueldo bruto anual": 24500},
    "Ceuta y Melilla": {"Edad media": 38, "Sueldo bruto anual":
21200},
}

# Función para calcular probabilidad de hipoteca
def calcular_probabilidad(perfil, comunidad1, comunidad2):
    try:
        edad = perfil["Edad media"]
        sueldo = perfil["Sueldo bruto anual"]

        if not comunidad1 or not comunidad2:
            return "Por favor, selecciona ambas comunidades
autónomas."

        # Filtrar datos por las comunidades seleccionadas
        datos_com1 = df[(df['CCAA'] == comunidad1) & (df['Año'] >=
2014) & (df['Año'] <= 2023)]
        datos_com2 = df[(df['CCAA'] == comunidad2) & (df['Año'] >=
2014) & (df['Año'] <= 2023)]

        if datos_com1.empty or datos_com2.empty:
            return "No se encontraron datos para una o ambas
comunidades seleccionadas."

        # Calcular medias de las columnas relevantes
        sueldo_com1 = datos_com1['Sueldo medio'].mean()
        sueldo_com2 = datos_com2['Sueldo medio'].mean()
        tasa_com1 = datos_com1['Tasa de desempleo de adultos
normalizado'].mean() if edad >= 26 else datos_com1['Tasa de
desempleo de jovenes normalizado'].mean()
        tasa_com2 = datos_com2['Tasa de desempleo de adultos
normalizado'].mean() if edad >= 26 else datos_com2['Tasa de
desempleo de jovenes normalizado'].mean()
        hipotecas_com1 = datos_com1['Hipotecas per capita
normalizado'].mean()
        hipotecas_com2 = datos_com2['Hipotecas per capita
normalizado'].mean()

        # Calcular Sedad
        if edad < 35:

```

```

        edad = 1
    elif 35 <= edad <= 50:
        edad = 1.5
    else:
        edad = 0.5

    # Cálculo del score con los nuevos pesos
    score_com1 = (0.55 * (sueldo / sueldo_com1)) + (0.15 * (1-
tasa_com1)) + (0.05 * 100 * hipotecas_com1) + (0.25 * edad)
    score_com2 = (0.55 * (sueldo / sueldo_com2)) + (0.15 * (1-
tasa_com2)) + (0.05 * 100 * hipotecas_com2) + (0.25 * edad)

    mejor_comunidad = comunidad1 if score_com1 > score_com2
else comunidad2
    return (f"Con el perfil seleccionado, es más probable
obtener una hipoteca en {mejor_comunidad}.\n"
           f"Score {comunidad1}: {score_com1:.2f}, Score
{comunidad2}: {score_com2:.2f}.")
except ValueError:
    return "Error en los datos proporcionados."

# Función para mostrar gráficos
def mostrar_graficos():
    variables = ['PIB per capita', 'Sueldo medio', 'Tasa de
desempleo']
    for var in variables:
        plt.figure(figsize=(8, 5))
        sns.scatterplot(x=df[var], y=df['Hipotecas per capita'],
hue=df['CCAA'], palette='Set2')
        plt.title(f"Relación entre {var} y Hipotecas per cápita")
        plt.xlabel(var)
        plt.ylabel("Hipotecas per cápita")
        plt.legend(bbox_to_anchor=(1.05, 1), loc='upper left')
        plt.show()

        plt.figure(figsize=(8, 5))
        sns.scatterplot(x=df[var], y=df['Importe de hipotecas per
capita'], hue=df['CCAA'], palette='Set2')
        plt.title(f"Relación entre {var} y Importe de hipotecas per
cápita")
        plt.xlabel(var)
        plt.ylabel("Importe de hipotecas per cápita")
        plt.legend(bbox_to_anchor=(1.05, 1), loc='upper left')
        plt.show()

    correlaciones = df[['Hipotecas per capita', 'Importe de
hipotecas per capita',
                        'PIB per capita', 'Sueldo medio', 'Tasa de
desempleo']].corr()

```

```

sns.heatmap(correlaciones, annot=True, cmap='coolwarm',
fmt=".2f")
plt.title("Mapa de calor de correlaciones")
plt.show()

# Crear interfaz en Google Colab con ipywidgets
def iniciar_aplicacion():
    perfil_dropdown = widgets.Dropdown(
        options=list(perfiles.keys()),
        description="Perfil:")
    comunidad1_dropdown = widgets.Dropdown(
        options=df['CCAA'].unique(),
        description="Comunidad 1:")
    comunidad2_dropdown = widgets.Dropdown(
        options=df['CCAA'].unique(),
        description="Comunidad 2:")
    graficos_button = widgets.Button(description="Mostrar
Gráficos")
    calcular_button = widgets.Button(description="Calcular")
    output = widgets.Output()

    def on_perfil_selected(change):
        with output:
            clear_output()
            perfil = perfiles[perfil_dropdown.value]
            print(f"Perfil seleccionado: Edad media: {perfil['Edad
media']}, Sueldo bruto anual: {perfil['Sueldo bruto anual']}")

    def on_graficos_clicked(b):
        with output:
            clear_output()
            mostrar_graficos()

    def on_calcular_clicked(b):
        with output:
            clear_output()
            perfil = perfiles[perfil_dropdown.value]
            resultado = calcular_probabilidad(perfil,
comunidad1_dropdown.value, comunidad2_dropdown.value)
            print(resultado)

    perfil_dropdown.observe(on_perfil_selected, names='value')

    calcular_button.on_click(on_calcular_clicked)
    graficos_button.on_click(on_graficos_clicked)

    display(widgets.VBox([
        widgets.Label("Selecciona un perfil y dos comunidades
autónomas para la comparación:"),

```

```

        perfil_dropdown,
        comunidad1_dropdown,
        comunidad2_dropdown,
        graficos_button,
        calcular_button,
        output
    ]))

# Ejecutar la función
iniciar_aplicacion()

# Importar bibliotecas
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
import ipywidgets as widgets
from IPython.display import display, clear_output
import numpy as np

# Cargar la base de datos desde el archivo de Excel
df = pd.read_excel('BBDD_TFG.xlsx')

# Crear las columnas necesarias si no están creadas
if 'Hipotecas per capita' not in df.columns:
    df['Hipotecas per capita'] = df['Número de hipotecas'] /
df['Población']

if 'Importe de hipotecas per capita' not in df.columns:
    df['Importe de hipotecas per capita'] = df['Importe de
hipotecas'] / df['Población']

# Normalizar columnas
def normalizar_columna(columna):
    return (columna - columna.min()) / (columna.max() -
columna.min())

df['Hipotecas per capita normalizado'] =
normalizar_columna(df['Hipotecas per capita'])
df['Sueldo medio normalizado'] = normalizar_columna(df['Sueldo
medio'])
df['Tasa de desempleo de jovenes normalizado'] = 1 -
normalizar_columna(df['Tasa de desempleo de jovenes'])
df['Tasa de desempleo de adultos normalizado'] = 1 -
normalizar_columna(df['Tasa de desempleo de adultos'])

# Perfiles medios por CCAA
perfiles = {
    "Andalucía": {"Edad media": 43, "Sueldo bruto anual": 21200},
    "Aragón": {"Edad media": 45, "Sueldo bruto anual": 25800},

```

```

    "Asturias": {"Edad media": 49, "Sueldo bruto anual": 24500},
    "Baleares": {"Edad media": 42, "Sueldo bruto anual": 22800},
    "Canarias": {"Edad media": 44, "Sueldo bruto anual": 21300},
    "Cantabria": {"Edad media": 46, "Sueldo bruto anual": 24200},
    "Castilla-La Mancha": {"Edad media": 44, "Sueldo bruto anual":
22400},
    "Castilla y León": {"Edad media": 48, "Sueldo bruto anual":
23700},
    "Cataluña": {"Edad media": 43, "Sueldo bruto anual": 28000},
    "Extremadura": {"Edad media": 46, "Sueldo bruto anual": 21000},
    "Galicia": {"Edad media": 48, "Sueldo bruto anual": 23500},
    "Madrid": {"Edad media": 43, "Sueldo bruto anual": 32000},
    "Murcia": {"Edad media": 42, "Sueldo bruto anual": 22300},
    "Navarra": {"Edad media": 44, "Sueldo bruto anual": 28500},
    "País Vasco": {"Edad media": 46, "Sueldo bruto anual": 30000},
    "La Rioja": {"Edad media": 45, "Sueldo bruto anual": 24000},
    "Valencia": {"Edad media": 44, "Sueldo bruto anual": 24500},
    "Ceuta y Melilla": {"Edad media": 38, "Sueldo bruto anual":
21200},
}

# Función para mostrar la matriz de scores
def mostrar_matriz_clicked(b):
    with output:
        clear_output()

        matrix_scores = pd.DataFrame(index=df['CCAA'].unique(),
columns=perfiles.keys())

        for comunidad in df['CCAA'].unique():
            for perfil_key, perfil_values in perfiles.items():
                edad = perfil_values["Edad media"]
                sueldo = perfil_values["Sueldo bruto anual"]

                datos_com = df[(df['CCAA'] == comunidad) &
(df['Año'] >= 2014) & (df['Año'] <= 2023)]

                if datos_com.empty:
                    matrix_scores.loc[comunidad, perfil_key] =
np.nan

                    continue

                sueldo_com = datos_com['Sueldo medio'].mean()
                tasa_com = datos_com['Tasa de desempleo de adultos
normalizado'].mean() if edad >= 26 else datos_com['Tasa de
desempleo de jovenes normalizado'].mean()
                hipotecas_com = datos_com['Hipotecas per capita
normalizado'].mean()

```

```

        if edad < 35:
            sedad = 1
        elif 35 <= edad <= 50:
            sedad = 1.5
        else:
            sedad = 0.5

        score_com = (0.55 * (sueldo / sueldo_com)) + (0.15
* (1 - tasa_com)) + (0.05 * 100 * hipotecas_com) + (0.25 * sedad)
        matrix_scores.loc[comunidad, perfil_key] =
score_com

    plt.figure(figsize=(12, 8))
    sns.heatmap(matrix_scores.astype(float), annot=True,
cmap='YlGnBu', fmt=".2f", linewidths=0.5, annot_kws={'size': 10})
    plt.title("Matriz de Scores para todos los perfiles y
comunidades autónomas")
    plt.xlabel("Perfiles")
    plt.ylabel("Comunidades Autónomas")
    plt.show()

# Interfaz simple con un solo botón
output = widgets.Output()
mostrar_matriz_button = widgets.Button(
    description="Mostrar matriz con todos los perfiles y CCAA",
    button_style='primary', # estilo visual (puede ser: 'success',
'info', 'warning', 'danger')
    layout=widgets.Layout(width='auto', height='60px',
padding='10px'),
    style={'font_size': '20px'}
)

mostrar_matriz_button.on_click(mostrar_matriz_clicked)

display(widgets.VBox([mostrar_matriz_button, output]))

```

Anexo 2: Archivo Excel con “BBDD_TFG.xlsx”