



Facultad de Ciencias Económicas y Empresariales

MODELO PREDICTIVO DE LOS PRECIOS DE LAS VIVIENDAS DE ALQUILER DE MADRID

Autor: Leticia de Wouters d'Oplinter Olalde
Clave: 201701059
Director: María de las Mercedes Barrachina

Resumen: Los precios de las viviendas en España han fluctuado abundantemente a lo largo de los años debido a diversos factores. Este año nos encontramos en máximos históricos en el precio de alquiler de una vivienda en España, una de las razones se debe a que estamos en una fase de crecimiento recuperando la crisis del COVID-19. El mercado inmobiliario es una de las opciones de inversión de algunas personas y por otro lado es la vía a la que recurren las personas para alquilar o comprar una casa. Es por ello, que es importante conocer los precios de las viviendas para saber cuál es el mejor momento para invertir o comprar una vivienda. Conocer los precios de antemano permite poder anticiparse y entender las posibles fluctuaciones de los precios. En este trabajo de investigación se ha realizado un modelo predictivo con un algoritmo para predecir el precio de las viviendas en alquiler que hay en Madrid.

Palabras clave: alquiler, precio, vivienda, mercado inmobiliario, oferta y demanda, hipoteca, tipos de interés, inflación, machine learning, performance predictiva, aprendizaje supervisado, modelo predictivo, regresión lineal, algoritmo y RStudio.

Abstract: Housing prices in Spain have fluctuated significantly over the years due to various factors. This year we have reached historic peaks in rental prices for houses in Spain, one contributing factor is the economic growth period recovering from the COVID-19 crisis. The real estate market is one appealing investment alternative for some individuals and additionally, it signifies how people rent or purchase property. Hence, it is important to be aware of housing prices in order to determine the most suitable time to invest or buy a house. Knowing the prices beforehand allows for anticipation and comprehension of possible price fluctuations. This thesis introduces a predictive model applying an algorithm to forecast rental house prices in Madrid.

Keywords: renting, price, housing, real estate market, supply and demand, mortgage, interest rates, inflation, machine learning, predictive performance, supervised learning, predictive model, linear regression, algorithm and RStudio.

ÍNDICE

1. Introducción	4
1.1 Objetivos.....	4
1.2 Metodología	5
2. Análisis del mercado inmobiliario español en los últimos años 20 años	5
3. Histórico de los precios de las viviendas de alquiler.....	11
3.1 Evolución del precio de las viviendas en España	11
3.2 Evolución de los precios de las viviendas por provincias en España	12
3.3 Características de los inquilinos de viviendas en España.....	13
3.4 Comparación de España con el resto de países.....	15
4. Factores macroeconómicos que afectan a los precios	17
5. Exploración de datos.....	21
5.1. Análisis de las variables	21
5.2. Análisis descriptivo	22
5.2.1. Resumen de estadísticos	22
5.2.2. Matriz de correlaciones	23
5.2.3. Distribución de las variables.....	25
5.3. Análisis del dataset.....	26
6. Modelo explicativo	28
6.1. Transformación de variables.....	28
6.2. Partición.....	29
7. Modelo algorítmico predictivo.....	30
7.1. Regresión lineal múltiple.....	30
7.2. Random forest.....	34
8. Resumen e interpretación de los resultados	36
9. Conclusión	38
10. Bibliografía	41

1. Introducción

El mercado inmobiliario de viviendas juega un papel importante en la economía de un país. Es un componente fundamental en la riqueza familiar y es muy importante en las decisiones de gasto y el bienestar financiero de las familias. Este principio se revela en la investigación del Banco de España, que establece que aproximadamente dos tercios de la riqueza total de las familias españolas están intrínsecamente ligadas a este mercado (Martínez & Maza, 2023). La importancia de comprender y anticipar las dinámicas de este sector es crucial para asegurar la estabilidad económica de las familias y, por lo tanto, del país en su conjunto.

El mercado inmobiliario es complejo debido a la variedad de factores que influyen en los precios de las viviendas. La variabilidad de los precios está directamente influenciada por cambios en la oferta y la demanda, condiciones macroeconómicas y eventos globales, lo que crea un entorno dinámico y desafiante (Montoriol, 2022). En este contexto, la implementación de un modelo predictivo se presenta como una estrategia crucial para no solo anticipar las tendencias futuras, sino también comprender cómo los diversos factores que dan forma al mercado de alquiler se relacionan entre sí.

En conclusión, en este estudio se desarrolla un modelo algorítmico para predecir los precios de las viviendas de alquiler en Madrid, el cual ayudará a comprender mejor la dinámica del mercado inmobiliario español. La aplicación de una técnica de aprendizaje supervisado será utilizada para prever y comprender las tendencias futuras y empoderar a los actores clave del mercado con herramientas analíticas que les permitan tomar decisiones estratégicas más informadas.

1.1 Objetivos

El objetivo de este trabajo es construir un modelo algorítmico, a través de una técnica de aprendizaje supervisado, para que según las distintas características de los pisos sea capaz de predecir los precios de las viviendas de alquiler de Madrid. Esto permitirá evaluar riesgos y ayudar a tomar decisiones. Otro de los objetivos es identificar las variables relevantes y cuáles son los factores que hacen que los precios fluctúen. Por último, otro

de los propósitos es analizar la performance predictiva del modelo para que sea lo más preciso posible.

1.2 Metodología

Para ello, lo primero que se hará es un análisis del mercado inmobiliario español en los últimos 20 años, para así comprender los factores que han sido importantes y han afectado a este mercado a lo largo del tiempo. A continuación, se profundizará en la evolución de los precios de alquiler para entender cómo han fluctuado los precios y el impacto que ha tenido una serie de cambios en los precios de las viviendas. También, se explicarán los factores macroeconómicos que afectan a los precios de las viviendas.

Después, se hará un modelo predictivo de aprendizaje supervisado, a partir de unas variables cualitativas y cuantitativas para desarrollar varios algoritmos capaces de predecir los precios de las viviendas de alquiler. La base de datos se centrará sólo en las viviendas de alquiler de Madrid, pero sería aplicable a mayor escala. Para la extracción de datos se utilizará una API creada en Octoparse, que es un software de extracción de datos web, para sacar los datos de la página web de Idealista. Para el desarrollo del modelo, se empleará Rstudio, que es un entorno de desarrollo integrado que utiliza el lenguaje de programación R. Una vez realizado el modelo se explicará cómo se ha creado y cuál es su performance predictiva, es decir, calcular los errores y la viabilidad de cada algoritmo, para poder elegir el que mejor se ajuste.

Por último, se interpretarán los resultados obtenidos y se analizarán, para así quedarnos con un solo modelo que sea el que menor error tenga, por lo tanto, el más fiable para la predicción de precios.

2. Análisis del mercado inmobiliario español en los últimos años 20 años

En los últimos años, el mercado inmobiliario español ha presentado fluctuaciones significativas, debido a cambios en la oferta y la demanda, como consecuencia de cambios económicos y políticos. A continuación, veremos más en detalle la evolución de este sector en las dos últimas décadas.

A principios de la década de los 2000 el mercado inmobiliario español era una inversión interesante para otros países europeos como puede ser Reino Unido, ya que los precios de las viviendas eran más baratos. En el año 2006, el precio de compra de una vivienda en España era de 2060 €/m² en comparación con Reino Unido que era de 2.130 €/m² (López-Rodríguez & Matea, 2019). Es por ello, que en España se empezaron a construir más inmuebles para responder a la nueva demanda. En consecuencia, esto hizo que entre los años 2000 y 2007 hubiese una elevada demanda, los precios empezaron a subir considerablemente, encontrándose en una situación económica favorable de crecimiento. Los compradores pedían hipotecas a los bancos para poder realizar sus inversiones, estos las concedían fácilmente sin estar seguros de que más tarde se las devolverán, por lo que era bastante accesible realizar la compra de un inmueble.

En el año 2003 empezaron a haber indicios de la formación de una burbuja en los precios de las viviendas en España. En 2005 se tenía un fácil acceso al crédito, obteniendo préstamos hipotecarios del 110% o 120% del precio de compra del inmueble y, como garantía, bastaba con utilizar la vivienda comprada. El aumento de la disponibilidad de préstamos hipotecarios con condiciones laxas no solo facilitó la adquisición de casas, sino que también contribuyó a la fragilidad del sistema financiero. La práctica de utilizar las mismas viviendas como garantía y conceder préstamos por encima del valor real de las propiedades creó un entorno propicio para una crisis financiera posterior (Arellano & Bentolia, 2009). La vulnerabilidad del sistema se agravó como resultado de una serie de factores, incluida la liberalización del mercado hipotecario y la falta de regulación adecuada.

La alta demanda de viviendas provocó un incremento en el número de viviendas construidas. Es por ello que, se invirtió en construcción y esto a su vez generó nuevos empleos. Esto favoreció la llegada de nuevos inmigrantes, aumentó la renta y favoreció el crecimiento del consumo privado a gran velocidad. Por otra parte, tuvo una repercusión positiva sobre el aumento del gasto de las familias y el crecimiento del PIB, ya que el incremento del precio de la vivienda provocó la generación de grandes plusvalías y los precios aumentaron debido a la posibilidad de acceder a una financiación bancaria de manera fácil.

El crecimiento de la construcción en España fue muy alto, del 5% al año entre 1996 y 2007. Entre 1998 y 2007 la zona urbanística española creció en 5.7 millones, casi el 30%.

A finales de 2007 la construcción suponía el 13,3% del empleo total, por encima del 6,7% de Alemania o del 8,5% del Reino Unido (Arellano & Bentolila, 2009).

El mejor año de la venta de viviendas en España fue en el año 2006, cuando se vendieron un total de 955,000 viviendas en todo el país. Por lo contrario, en el año 2013 fue cuando se vendieron el menor número de viviendas. Esto significa que desde el 2006 hasta el 2013 las ventas bajaron un 70% y los precios de las viviendas se derrumbaron.

En el año 2008 estalló la burbuja inmobiliaria, cayó la demanda de manera significativa y a su vez los precios, alcanzando su punto máximo. La principal causa de esta crisis fue por la incapacidad del mercado para absorber la gran oferta de las viviendas construidas. Para entender mejor la bajada de los precios es importante aclarar que una burbuja especulativa se produce cuando los precios suben de forma significativa, de manera anormal, prolongada e incontrolada debido a la especulación. La especulación consiste en comprar a un precio y venderlo en un tiempo a un precio mayor, para obtener un beneficio en la inversión. Es por ello, por lo que con el paso de los años el precio sube desmesuradamente hasta que se produce un estallido y se desploma. En ese momento es cuando se empiezan a vender los bienes y hay pocos compradores dispuestos a adquirirlo, por lo que hay un exceso de oferta y el precio baja de forma considerada quedándose en niveles muy bajos (Bernardos, 2009).

El ciclo inmobiliario español experimentó un cambio debido a factores internos y externos. En primer lugar, la burbuja inmobiliaria fue consecuencia del elevado incremento de la demanda de viviendas. Aunque, la oferta aumentó de forma considerable, no lo hizo en la proporción suficiente, ni con la rapidez necesaria para absorber la nueva demanda sin generar un aumento en los precios. En segundo lugar, se debe a la falta de liquidez del sistema financiero, causada por la crisis de las hipotecas subprime en Estados Unidos, por el deterioro interno de la economía española, la falta de financiación y el agotamiento del modelo de crecimiento basado en la construcción.

En 2014, la vivienda había perdido un cuarto de su valor desde máximos en 2006, cuando el precio era de 2060 €/m², cayendo su precio por metro cuadrado un 25% quedándose en 1.564 €/m² (Idealista, 2014). Desde la crisis de 2008 los precios empezaron a bajar constantemente hasta el año 2014. Una de las causas de esta bajada fue el aumento del

paro debido a la crisis que hizo que las personas no quisiesen arriesgar a comprar una vivienda o cambiar de residencia.

Después de la crisis, se tomaron varias medidas para estabilizar el mercado inmobiliario en España. El objetivo de las reformas financieras y cambios en las políticas de vivienda fue recuperar la confianza de los inversores e impulsar la recuperación económica. Sin embargo, la recuperación fue gradual y el mercado inmobiliario continuó lidiando con problemas, como administrar la sobreoferta de casas construidas durante el auge previo. En respuesta a los desafíos y lecciones aprendidas de la crisis económica de 2008, el entorno regulatorio del mercado inmobiliario español experimentó transformaciones significativas. El objetivo de la implementación de regulaciones fue aumentar la estabilidad y sostenibilidad del sector inmobiliario. Por ejemplo, la Ley 5/2019, que regula los contratos de crédito inmobiliario, mejoró la protección de los consumidores que reciben préstamos hipotecarios. Esta ley promueve la transparencia y reduce la disparidad de información entre prestamistas y prestatarios al establecer claramente los derechos y obligaciones de las partes involucradas (BOE, 2019).

El Gobierno español implementó regulaciones como el Código Técnico de la Edificación (CTE) con el objetivo de mejorar la eficiencia energética en el sector inmobiliario. El CTE fomenta el uso de tecnologías y prácticas que reduzcan el impacto ambiental de los nuevos desarrollos inmobiliarios, estableciendo requisitos mínimos de eficiencia energética y sostenibilidad en la construcción de edificaciones (CTE, 2019).

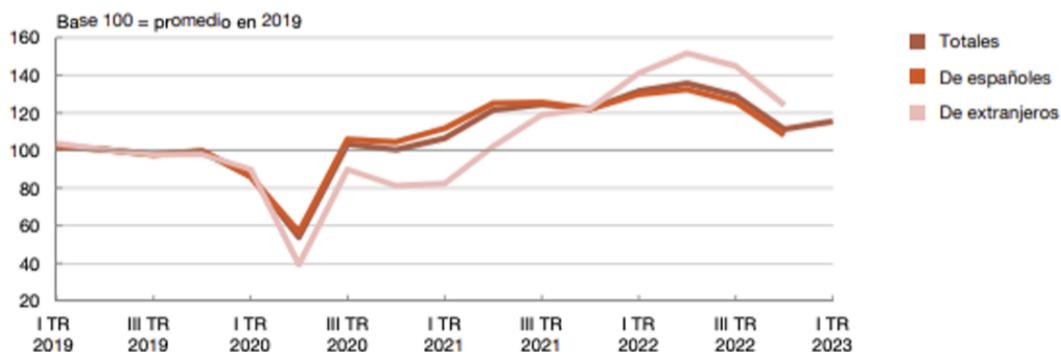
Estas iniciativas de regulación muestran el nivel de compromiso de España en fortalecer la transparencia, la protección del consumidor y la sostenibilidad en el sector, contribuyendo a la creación de un mercado más sólido y equitativo.

Durante el periodo de recesión hubo una fuerte bajada de precios que duró hasta el 2014, y a partir de ahí volvió a crecer hasta el COVID. El año 2014 fue el punto de inflexión del sector inmobiliario. A partir de ahí las ventas volvieron a crecer, se volvieron a demandar hipotecas, los precios volvieron a incrementar en algunas provincias, la vivienda se consolidó como inversión rentable y el alquiler de viviendas volvió a aumentar. Es en ese año, 2014, en el que se pone fin a la recesión inmobiliaria y empieza la época de crecimiento (Alves & San Juan, 2018).

A principios de 2020 apareció el COVID-19, la pandemia que paralizó el mundo. Debido a la escasez de materias primas vivida en 2021, como puede ser la falta de mano de obra cualificada, los costes encarecieron, y esto afectó directamente a los plazos de entrega, que se prorrogaron y a la oferta, lo que supuso una subida de precios (Alves & San Juan, 2021).

La sensibilidad del mercado inmobiliario español ha sido demostrada por los acontecimientos globales que tienen un impacto en la economía mundial. La reciente pandemia de COVID-19, tuvo un impacto significativo en el sector inmobiliario en España y a nivel mundial, es un ejemplo destacado. La crisis sanitaria alteró las transacciones inmobiliarias, reduciendo considerablemente la demanda de viviendas y cambiando la forma en que se realizan las transacciones inmobiliarias, como el aumento de las visitas virtuales a las propiedades (Knight Frank, 2021). Adicionalmente, la incertidumbre económica global tiene un impacto directo en la confianza de los inversores y consumidores, influyendo en las decisiones de compra y venta de viviendas.

Figura 1: Evolución de la demanda de vivienda en España 2019-2023



Tras la crisis vivida durante y después de la pandemia, el mercado inmobiliario volvió a cobrar importancia entre los inversores. Este crecimiento se debe a que el criterio del consumidor cambió y se empezaron a demandar viviendas más grandes con mayores espacios abiertos. Esta demanda junto con los tipos de interés muy bajos hizo que se notasen los cambios en el mercado inmobiliario.

En el año de la pandemia y durante el confinamiento las ventas de las viviendas en España se desplomaron. Aunque, más adelante cambió el tipo de demanda tradicional y pasó a demandarse casas independientes. A día de hoy este tipo de demanda se ha reducido, ya que no hay mucha oferta, debido a que históricamente se ha demandado más el piso tradicional.

La inversión en el mercado inmobiliario ha sido notable en este último año. Desde la pandemia se han cambiado los hábitos de los consumidores a viviendas más grandes con más terreno exterior y esto ha impulsado la demanda de viviendas en España.

En cuanto las perspectivas futuras del mercado inmobiliario español, muchos factores internos y externos tienen un gran impacto. Las políticas gubernamentales, especialmente las relacionadas con la vivienda, la planificación urbana y la sostenibilidad, tendrán un impacto significativo en la configuración del mercado a nivel interno. El desarrollo sostenible del sector dependerá de la implementación de medidas que fomenten la accesibilidad a la vivienda, promuevan la eficiencia energética y aborden la creciente demanda. A nivel externo, el mercado inmobiliario español se verá afectado por los siguientes tres puntos: los cambios en la economía mundial, las tasas de interés y la estabilidad geopolítica. Es importante tener en cuenta los avances de la sociedad como la transición hacia energías más sostenibles, cambios en la movilidad laboral y la resiliencia del mercado frente a posibles crisis económicas.

En conclusión, durante las últimas décadas, el mercado inmobiliario español ha atravesado diversas fases, desde un período de prosperidad y expansión acelerada hasta la crisis económica de 2008, la cual provocó cambios significativos en el marco regulatorio. La capacidad del mercado para enfrentar desafíos se ha demostrado a través de la recuperación y las adaptaciones a eventos globales, como la pandemia de COVID-19. Para garantizar un crecimiento equitativo y sostenible en el futuro, será necesaria una planificación urbana sostenible, transparencia en las transacciones inmobiliarias y consideración de factores globales. La cooperación entre los sectores público y privado, así como la capacidad de adaptarse a las tendencias globales, serán fundamentales para mantener el éxito del mercado inmobiliario español en las próximas décadas.

El análisis destaca la complejidad y la variedad del mercado inmobiliario español durante los últimos veinte años, mostrando la influencia de factores económicos y políticos en su desarrollo. Como se discutirá en las secciones posteriores de esta tesis, la comprensión de estos eventos históricos y los cambios en los precios proporciona un contexto útil para la creación de modelos predictivos.

3. Histórico de los precios de las viviendas de alquiler

Como hemos visto en el apartado anterior, por distintas razones, el precio de la vivienda en España ha fluctuado significativamente debido a diversas variables económicas y financieras. En esta sección, profundizaremos en la evolución detallada de los precios de las viviendas desde 2007 hasta la actualidad, ofreciendo una comprensión más profunda de los factores que han influido en estos cambios.

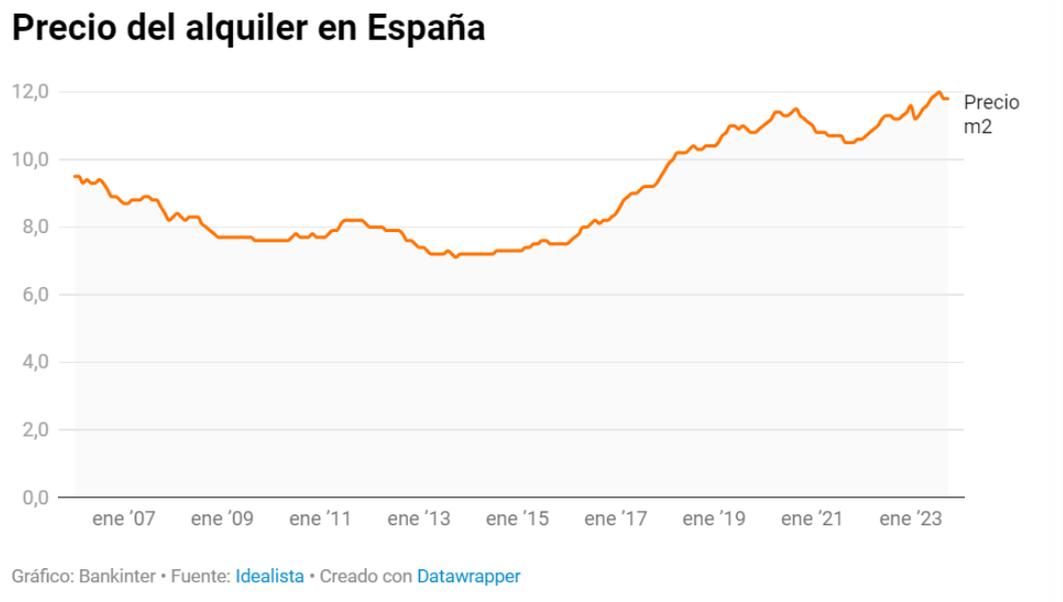
3.1 Evolución del precio de las viviendas en España

En la Figura 2 podemos observar que en el año 2007 el precio de alquiler de una vivienda en España era de 8,6€/m² (Idealista, 2007). A partir de ahí el precio siguió bajando hasta alcanzar los mínimos en el año 2014 con un precio de 7,2€/m² (Idealista, 2014). Esta disminución se asocia con diversos factores, incluyendo la secuela de la crisis financiera. Desde entonces el precio de alquiler de vivienda en España empezó a subir hasta día de hoy, teniendo un único punto de inflexión en el año 2021 a causa del COVID, donde el precio bajó a 10,4€/m². Como podemos observar en la gráfica, este año nos encontramos en máximos históricos desde el 2007, con un precio de 11,8€/m², ya que estamos en un periodo de crecimiento económico (Idealista, 2023).

El portal Idealista, reconocido por sus análisis exhaustivos del mercado inmobiliario, ha sido una fuente valiosa para comprender estos movimientos. Sin embargo, es necesario considerar otras perspectivas para obtener una visión completa del panorama. Según el Instituto Nacional de Estadística (INE), la recuperación económica general y la estabilización del mercado laboral han influido en el aumento de los precios de alquiler (INE, 2023). Además, el Banco de España ha señalado la importancia de las políticas

monetarias y fiscales en la dinámica de los precios de la vivienda en España (Mayordomo Tello, 2019).

Figura 2: Evolución de los precios de las viviendas de alquiler en España 2007-2023



3.2. Evolución de los precios de las viviendas por provincias en España

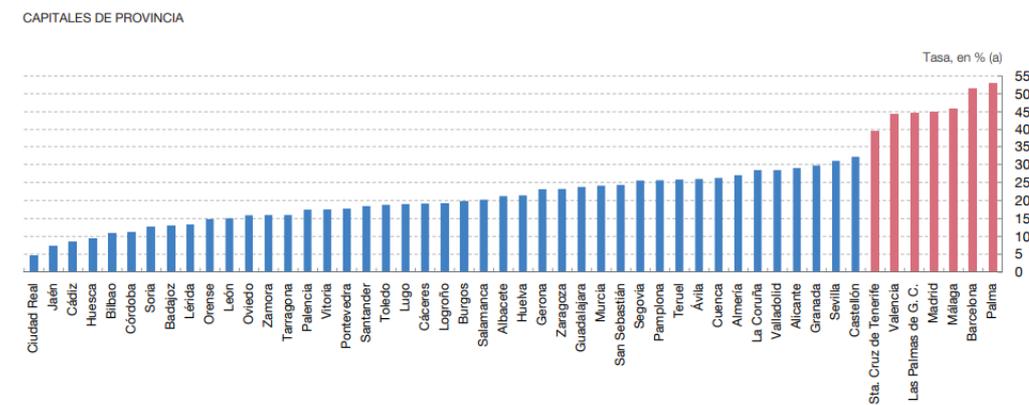
Los precios de alquiler en España pueden variar significativamente según la provincia debido a algunos factores geográficos, demográficos y económicos. Estas diferencias se deben principalmente a la oferta y a la demanda que haya en esa provincia, al tamaño, a la ubicación: si se encuentra en zona interior o zona costera, al crecimiento económico y entre otras, si es una provincia turística (Martínez & Maza, 2023). La provincia más cara para alquilar un piso en España en octubre de este año es Barcelona con un precio de 20€/m² y la menos cara es Jaén con un precio de 5,6 €/m² (Idealista, 2023).

En cuanto a Madrid, también existe una variación de precios en función del distrito, debido a la ubicación, la oferta, la demanda, la seguridad, la calidad de vida, las zonas comerciales, las oportunidades laborales y el desarrollo urbanístico. Los distritos más caros de viviendas de alquiler dentro de Madrid en octubre de este año son: Salamanca, Centro y Chamberí con un precio de más de 20€/m². Los más baratos son Vallecas, Vicálvaro y Villaverde con un precio alrededor de 12€/m² (Idealista, 2023). La diferencia

entre ambos es considerable, ya que se trata casi del doble del precio dentro de Madrid. Por lo tanto, puede haber grandes diferencias de precio dentro de una misma provincia, pero depende de su tamaño, ya que Madrid es una ciudad bastante grande. Por ejemplo, Toledo es una ciudad más pequeña y la diferencia entre el distrito más caro y el más barato en octubre de este año es de 2,3€/m², siendo el más caro Toledo con un precio de 8,3€/m² y el más barato Talavera de la Reina con un precio de 6€/m² (Idealista, 2023).

Como hemos visto anteriormente los precios han ido creciendo en los últimos años, pero no ha sido igual para todas las provincias. En la figura 3, podemos observar el crecimiento acumulado del precio medio de oferta de alquiler por provincia en España entre los años 2014 y 2019. Las provincias con mayor tasa de crecimiento son las capitales con mayor población: Madrid, Barcelona y Valencia. También, tienen una tasa alta las provincias con elevada actividad turística como Málaga y las Islas Canarias. Las provincias con menor crecimiento acumulado son las que se encuentran en el interior de España (López-Rodríguez & Matea, 2019).

Figura 3: Crecimiento acumulado del precio medio de la oferta del alquiler 2014-2019



FUENTES: Idealista y Banco de España.

a El crecimiento acumulado se calcula como la tasa de variación entre el precio medio de diciembre de 2013 y el de mayo de 2019.

3.3 Características de los inquilinos de viviendas en España

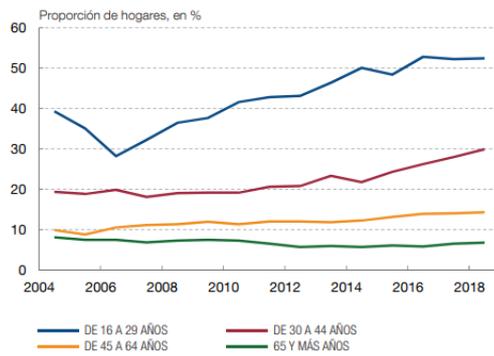
A continuación, analizaremos las características propias de las personas que tienden a alquilar viviendas en vez de comprarlas para entender mejor el perfil de los inquilinos. Según un estudio que realizó el Banco de España en el año 2019, la edad de las personas

es un factor determinante que influye a la hora de alquilar un piso (López-Rodríguez & Matea, 2019). Podemos observar en la figura 4 que vemos abajo, que en 2018 más del 50% de los pisos alquilados en Madrid fueron de jóvenes de entre 16 y 29 años. Esto tiene una relación directa a que esos jóvenes tienen recursos financieros limitados. El alquiler generalmente implica una carga financiera menor en términos de costes iniciales y mensuales en comparación con la compra. A su vez, es una opción más viable en términos de accesibilidad financiera dentro del mercado competitivo. Otra de las razones, es que es muy útil para las jóvenes que estén en una etapa de transición y esto les da más flexibilidad, ya que les permite cambiar de ubicación más fácilmente, sin tener que comprometerse. Desde el año 2006 hasta el 2018 el porcentaje de jóvenes ha aumentado considerablemente, comenzando con un 30% llegando a superar el 50%. Como vemos en la gráfica, según va aumentando la edad, el porcentaje de personas que alquilan viviendas disminuye. En el año 2018, sólo un 7% de las personas que viven en un piso de alquiler tienen más de 65 años, esta cifra se ha mantenido constante desde 2004 (López-Rodríguez & Matea, 2019).

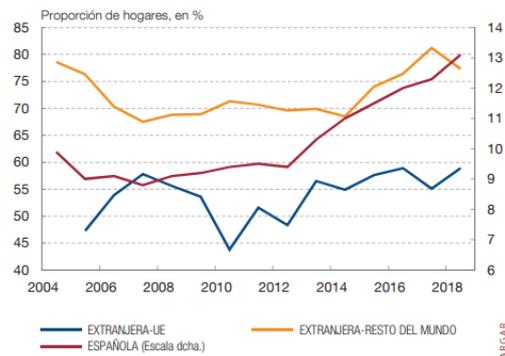
Otro de los factores característicos a tener en cuenta para poder analizar a las personas que viven en pisos de alquiler es su nacionalidad. Como vemos en la parte derecha de la figura 4 de abajo en el año 2018 cerca del 80% de las personas que viven en un piso de alquiler son extranjeras. Una de las razones, es que acaban de llegar a un país nuevo y todavía no están invirtiendo en comprar una vivienda, ya que no conocen el mercado inmobiliario del país y se sienten más cómodos alquilando que comprando. A veces, el proceso de compra puede ser complicado debido a las regulaciones y leyes locales. Otra de las razones, es que algunas personas extranjeras se mudan temporalmente, ya sea por trabajo, estudios u otras razones y prefieren alquilar para tener una mayor flexibilidad, ya que tienen incertidumbre sobre su permanencia. Esta opción les parece más conveniente y así no tienen que comprometerse a largo plazo. Los españoles son los que tienen el porcentaje más bajo, con un 13% en 2018, que ha ido creciendo desde 2005, ya que contaban sólo con un 9%. Estos porcentajes son mucho más bajos, debido a que algunos españoles han decidido hacer una inversión en este sector y tienen más clara su permanencia (López-Rodríguez & Matea, 2019).

Figura 4: Evolución del alquiler residencial en España 2004-2018

1 ALQUILER RESIDENCIAL, POR GRUPOS DE EDAD



2 ALQUILER RESIDENCIAL, POR NACIONALIDAD



DESCARGAR

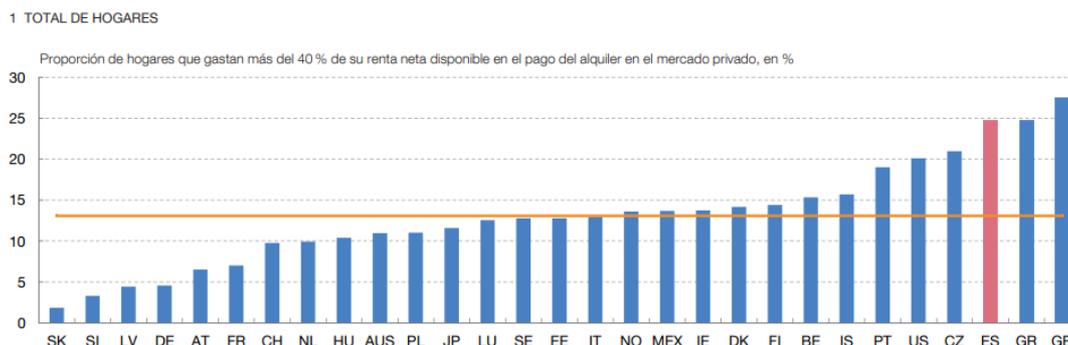
Estas características descritas anteriormente sirven para conocer mejor la demanda de las viviendas y el tipo de persona que prefiere vivir en un piso de alquiler, para así poder adaptar mejor los pisos de alquiler a las necesidades de los inquilinos y poder satisfacerlas (Trecet, 2021).

En 2022 la oferta total que hay en España de pisos de alquiler es de 4,2 millones, que se encuentra por debajo de la demanda actual que es de 5,4 millones. Según la agencia inmobiliaria Savills (2022), se necesitan 1,2 millones de viviendas “build to rent” para frenar el desequilibrio entre la demanda y la oferta.

3.4 Comparación de España con el resto de países

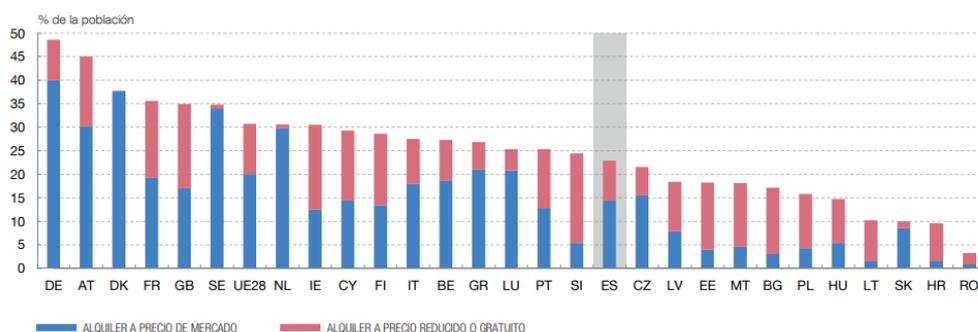
Una vez hemos analizado los precios de España y de sus provincias, así como el perfil de los inquilinos de viviendas españolas, es importante contextualizarlo a nivel mundial. En el año 2014 España presentó, entre las economías de la OCDE (Organización para la Cooperación y el Desarrollo Económico), una de las mayores proporciones de hogares en el que el gasto de vivienda de alquiler supone una sobrecarga en relación con sus ingresos neto. El 25% de los hogares gastan más del 40% de su renta neta en el pago del alquiler de sus viviendas. España se encuentra en el tercer puesto después de Reino Unido y Grecia (López-Rodríguez & Matea, 2019).

Figura 5: Índice de sobrecarga del gasto en vivienda de alquiler en los países de la OCDE en 2014



Para entender mejor la posición que tiene España frente a otros países de la UE analizaremos la figura 6 que se muestra a continuación. En 2017 España se encontraba por debajo de la media de la UE en la proporción de residentes que habitan en viviendas de alquiler y cuenta tan sólo con un 24% de la población. Esto significa que la mayor parte de los residentes viven en pisos comprados, residencias o similares (López-Rodríguez & Matea, 2019).

Figura 6: Población residente en vivienda de alquiler en los países de la UE en 2017



FUENTE: Encuesta Europea de Ingresos y Condiciones de Vida (EU-SILC) [Eurostat (2018)].



En conclusión, el precio de alquiler de una vivienda en España ha ido cambiando a lo largo de los años, encontrándose hoy en día en máximos históricos (Polanco, 2023). La evolución de los precios se ve afectada por diferentes características como puede ser la oferta y la demanda, políticas gubernamentales, cambios en los tipos de interés, inversión extranjera y ciclos económicos. Además, dependiendo de la provincia estos precios

también varían en función de otros factores mencionados anteriormente. En estos últimos años la oferta de nuevas viviendas ha sido mayor que la demanda, y a su vez ha habido escasez en el mercado de segunda mano en las ciudades principales como Madrid y Barcelona. Es por ello que, se ha mantenido una demanda elevada y los precios han ido creciendo. Se prevé que para el 2024 haya una corrección en los precios de las viviendas a nivel nacional. Asimismo, el Banco Central Europeo ha insinuado la posibilidad de aumentar los tipos de interés. La consecuencia de estas subidas sería que las hipotecas fuesen más caras, por lo que la demanda bajaría y a su vez los precios, ya que sería más costoso obtener un préstamo y la inversión inmobiliaria dejaría de ser tan atractiva para algunas personas (Idealista, 2023)

4. Factores macroeconómicos que afectan a los precios

Después de haber visto la fluctuación de los precios a lo largo de los años, vamos a analizar los principales factores macroeconómicos que tienen una influencia significativa en el mercado inmobiliario y hacen que los precios varíen.

En primer lugar, uno de los factores macroeconómicos importante es la inflación, no sólo su efecto en el poder adquisitivo de las personas, reduciendo la demanda, sino también el precio y la escasez de los materiales utilizados para la construcción de las viviendas. La escasez de los materiales empezó con la pandemia del COVID-19 y se prolongó debido al conflicto ruso-ucraniano por el cual hubo que modificar la logística, lo cual hizo que el precio de las materias prima se disparase, retrasando la construcción de viviendas y aumentó el costo general.

El incremento de los precios de construcción también se debe al incremento del importe de edificación residencial, que según ACR se incrementó en 23,5% a finales de 2021. (ACR, 2022). Una vez estábamos recuperándonos de la pandemia apareció el conflicto entre Rusia y Ucrania, el cual complicó las rutas de distribución europeas. La relocalización de recursos de los gobiernos para financiar a sus aliados, y el parón de recolección de algunas materias primas fue lo que desencadenó una subida de precios considerable. Las materias primas más afectadas por este conflicto fueron el aluminio y la arcilla. Desde que empezó el conflicto, el precio de la madera ya se paga tres veces

menos. Por comparativa en 2022 los precios de los materiales de construcción eran un 38% más que en 2020 (ACR, 2022).

Según un estudio de CBRE la obra que más sufre es la no residencial, ya que los materiales que usan son los que más han subido de precio. La media de retrasos en este tipo de construcciones es de entre 2 a 3 meses al tiempo inicial establecido. En cuanto el incremento medio del precio se sitúa entre 20-25%, en espacios logísticos se situaría entre el 8-10%. En cuanto a las obras residenciales prevé que, al aumentar los costes de producción de viviendas, se aumenta el precio de venta, en torno al 3-5%, haciendo que la oferta de las obras residenciales disminuya (CBRE, 2022).

No solo han aumentado los costes de las materias primas, sino también otros gastos que afecta directamente a la construcción, como el precio de la energía que alcanzó uno de sus picos en 2023, también el precio del combustible utilizado para transportar los materiales. Volviendo a las materias primas podemos ver que algunos materiales de la construcción han incrementado su precio considerablemente desde 2019 a 2022 como: acero (83,70%), madera (27,70%), hormigón (23,30%) y cemento (19,6%) (Banegas, 2022).

Según CBRE (2022) se prevé alcanzar el equilibrio de los precios en torno a 2025, esto es algo positivo para el mercado español, ya que promoverá la inversión extranjera en construcción de viviendas. También, prevé que la disminución de la presión inflacionista en materias primas y la resolución de problemas en la cadena de suministro y producción de materiales, que surgieron a raíz de la pandemia y se intensificaron con el conflicto en Ucrania, reducirán el aumento de costes en la construcción en los años próximos. En los últimos meses se ha observado una disminución en los precios de los metales industriales, y se prevé que esta tendencia a la baja se mantenga.

Otro de los factores macroeconómicos que afectan directamente a los precios de las viviendas son los tipos de interés que son clave en la dinámica del mercado inmobiliario debido a su impacto directo en la accesibilidad financiera para comprar viviendas. En aquellos momentos en los que las tasas de interés son bajas, los préstamos son más accesibles, lo que motiva a los consumidores a recurrir a la financiación hipotecaria, y como resultado aumenta la demanda de propiedades. Varios estudios han documentado

este fenómeno, incluido el de Reichert (1990), que enfatiza cómo las bajas tasas de interés impulsan el crecimiento del sector inmobiliario al facilitar la adquisición de propiedades.

No obstante, la situación cambia con el incremento de los tipos de interés. Según Harris (1989), un aumento en las tasas puede reducir drásticamente la actividad inmobiliaria al aumentar los costes de los préstamos hipotecarios. Dado que la accesibilidad financiera disminuye, este fenómeno puede tener un impacto tanto en la demanda como en los precios de las viviendas. La relación inversa que existe entre las tasas de interés y la actividad del mercado inmobiliario subraya la necesidad de observar cuidadosamente los tipos para entender y anticipar fluctuaciones en los precios de las viviendas.

Es importante resaltar que los precios en el mercado inmobiliario también se determinan principalmente por el principio fundamental de oferta y demanda. Los valores de las propiedades pueden verse significativamente afectados por la dinámica interacción entre la oferta de viviendas disponibles y la demanda de compradores potenciales. Investigaciones, como el estudio de de la Paz (2006), muestran cómo un exceso de oferta puede resultar en una disminución de precios mientras que una fuerte demanda puede generar aumentos considerables.

Adicionalmente, según un estudio que se presentó en la "International Conference on Management Science and Engineering" (Wan et al., 2009), los cambios en la estructura familiar, las preferencias de los compradores y los factores demográficos pueden afectar la demanda. Esto enfatiza la complejidad de la dinámica de oferta y demanda en el mercado inmobiliario.

Por último, uno de los factores macroeconómicos más recientes que ha afectado al mercado inmobiliario es la ley de vivienda 2023, que es una legislación diseñada para desarrollar el derecho a una vivienda digna. Esta ley tiene como objetivo principal ayudar a aquellos colectivos que se enfrentan a mayores dificultades para acceder a una vivienda. Fue aprobada por el Congreso y el Senado, entró en vigor el 26 de mayo de 2023. En esta nueva ley podemos identificar nuevos cambios que benefician tanto al inquilino como al propietario, aunque esté más centrada a otorgar una vivienda digna al inquilino. A partir de ahora será el propietario el que se encargará de pagar los honorarios a la agencia

inmobiliaria, haciendo que el precio de la vivienda disminuya para el inquilino (BOE, 2023).

También, podemos observar que se introduce el concepto de áreas de alta tensión del mercado inmobiliario, son áreas en las que los precios de las viviendas son tan elevados, que supone un problema significativo para que las familias puedan acceder a ellas. Las comunidades autónomas se encargarán de nombrar esas zonas, para declarar dicha zona tienen que cumplir uno de los dos requisitos: que el precio de la vivienda haya aumentado más de 3 puntos el valor del IPC en los últimos 5 años o que el precio medio supere el 30% de los ingresos medios de la zona.

Esta ley también viene con beneficios interesantes para los propietarios que posean una vivienda en una zona tensionada, mediante incentivos que consisten en una bonificación del IRPF. Se otorgará una bonificación del 50% a todos los propietarios que alquilen su vivienda, un 60% si la vivienda ha sido objeto de una actuación de rehabilitación, un 70% si se alquila por primera vez a jóvenes de entre 18-35 años y un 90% en casos donde la renta inicial se haya reducido más de un 5% en comparación con la última renta del contrato de arrendamiento anterior, aplicando la cláusula de actualización anual del contrato previo (BOE, 2023).

Por otro lado, se modificará la definición de gran propietario. Un gran propietario se consideraba una persona que poseía 10 inmuebles o más, o una superficie construida que superase los 1.500 m². Ahora las comunidades autónomas pueden disminuir el número de viviendas a 5, si se encuentra en una zona que haya sido declarada tensionada (BOE, 2023).

Para terminar, se ha establecido un límite a la actualización del alquiler para controlar la inflación, que aumentara del 2% al 3% para 2024. Consecuentemente buscan crear un nuevo índice de referencia de precios de alquiler que sustituya al IPC, que tendrá efecto a partir de 2025. El índice estará limitado a 2% en 2023, 3% en 2024 y en 2025 todavía depende del nuevo índice de actualización (BOE, 2023).

5. Exploración de datos

Anteriormente, hemos visto el impacto significativo que puede tener la valoración de las viviendas en la economía de un país y también a nivel mundial. Es por ello, que a continuación, se elaboran y explican varios modelos algorítmicos para predecir los precios de las viviendas de alquiler de Madrid. Los modelos algorítmicos que se han creado son de predicción, utilizando la técnica de aprendizaje supervisado, que consiste en separar el data set en dos partes: training set y test set, para entrenar al algoritmo a predecir los precios. Con la parte del training set se entrena al algoritmo y con la parte del test set se comprueba cómo funciona el algoritmo y cuanto es el error para cada observación entre el valor real y el valor que ha precedido el algoritmo (IBM, n.d.).

5.1. Análisis de las variables

En esta sección, analizaremos las variables del dataset que se han empleado para la creación del modelo algorítmico de predicción. El dataset se ha extraído del siguiente link: <https://www.idealista.com/alquiler-viviendas/madrid-provincia/> de la página web de idealista, la agencia inmobiliaria española, ya que es la que proporciona los datos más precisos. Los datos han sido extraídos en el mes de noviembre de este año y son de los pisos en alquiler de Madrid. La extracción de los datos se ha hecho a través de una API desarrollada con el programa Octoparse que es una herramienta de web scarping. Dado que la página web tiene restricciones para que no se haga un uso masivo de sus datos contamos con 8.118 observaciones. Se utilizó la API para la descarga de datos en otras páginas web de inmobiliarias como puede ser Fotocasa, Gilmar o Rentalia, pero no todas proporcionaban el mismo tipo de variables, por lo que los datos no eran homogéneos y podían existir duplicados. Es por ello que, para realizar los modelos algorítmicos de predicción se ha empleado únicamente el dataset extraído de la inmobiliaria idealista.

El dataset cuenta con 7 variables de carácter numérico. La variable explicativa del modelo es la variable “precio” que es de carácter numérico está expresada en euros e indica el precio de la vivienda. Es aquella que va a ser entrenada por el modelo, para ser capaz de predecir los precios de las nuevas observaciones. Las demás variables son las variables independientes del modelo. A continuación, se muestran el resto de las variables:

- Número de habitaciones: variable numérica que indica el número de habitaciones que tiene cada vivienda.
- Número de baños: variable numérica que indica el número de baños que tiene cada vivienda
- Número de m²: variable numérica que indica el número de metros cuadrados que tiene cada vivienda
- Planta o bajo: variable numérica dicotómica que toma el valor 0 si la vivienda se encuentra en un bajo o valor 1 si se encuentra en una planta.
- Garaje: variable numérica dicotómica que toma el valor 0 si la vivienda no tiene garaje o valor 1 en el caso de que la tenga
- Ascensor: variable numérica dicotómica que toma el valor 0 si la vivienda no tiene ascensor y valor 1 en el caso de que lo tenga.

5.2. Análisis descriptivo

En este apartado se hará un análisis de las variables en Rstudio para entender la relación que hay entre ellas, su comportamiento y el tipo de variables.

5.2.1. Resumen de estadísticos

Para obtener un resumen de todas las variables se ha utilizado la función summary en el que aparece el mínimo valor que puede tomar esa variable, el máximo, el primer cuartil, la mediana, la media y el tercer cuartil (Llinás, 2021). Esto ayuda a entender cuáles son los valores que toman las observaciones que tenemos en el dataset.

Summary(datos)

Precio	baño	m2	numero de hab	planta o bajo
Min. : 350	Min. : 0.000	Min. : 12.0	Min. : 0.000	Min. :0.0000
1st Qu.: 1100	1st Qu.: 1.000	1st Qu.: 60.0	1st Qu.: 1.000	1st Qu.:1.0000
Median : 1540	Median : 2.000	Median : 84.0	Median : 2.000	Median :1.0000
Mean : 2270	Mean : 1.841	Mean : 119.8	Mean : 2.242	Mean :0.9069
3rd Qu.: 2764	3rd Qu.: 2.000	3rd Qu.: 131.0	3rd Qu.: 3.000	3rd Qu.:1.0000
Max. :30000	Max. :11.000	Max. :1739.0	Max. :15.000	Max. :1.0000
ascensor	Garaje			
Min. :0.0000	Min. :0.000			

1st Qu.:1.0000	1st Qu.:0.000
Median :1.0000	Median :0.000
Mean :0.7562	Mean :0.365
3rd Qu.:1.0000	3rd Qu.:1.000
Max. :1.0000	Max. :1.000

Al observar esta tabla podemos comprobar que el precio más bajo de las viviendas observadas es 350 y el más alto 30.000, aunque la mediana se encuentra en 1.540, es decir, el 50% de las viviendas tienen un precio menor a 1.540. En cuanto a la media vemos que está en 2.270. La variable baño toma como valor más alto el 11 y como mínimo 0, la media de baños por vivienda es de 1.84, por lo que la mayoría de pisos tienen alrededor de dos baños. El piso con menos metros cuadrados tiene 12m² y el que más tiene 1.739m², la media es 84m², por lo que el 50% de la muestra tiene entre 12 y 84m². Haciendo referencia al número de habitaciones, hay pisos con 0 habitaciones, esto significa que son estudios y el máximo está en 15, la media está en 2 habitaciones por vivienda. En cuanto a la variable planta o bajo es dicotómica, si observamos la media está en 0,906 muy cerca del 1, esto quiere decir que la mayor parte de las viviendas no son bajos, sino que están en alguna planta. En el caso de la variable ascensor la media está en 0,756 por lo que la mayoría de las viviendas tienen ascensor. Por último, la media de la variable garaje es 0,365 esto significa que hay pocas viviendas que tengan garaje.

5.2.2. Matriz de correlaciones

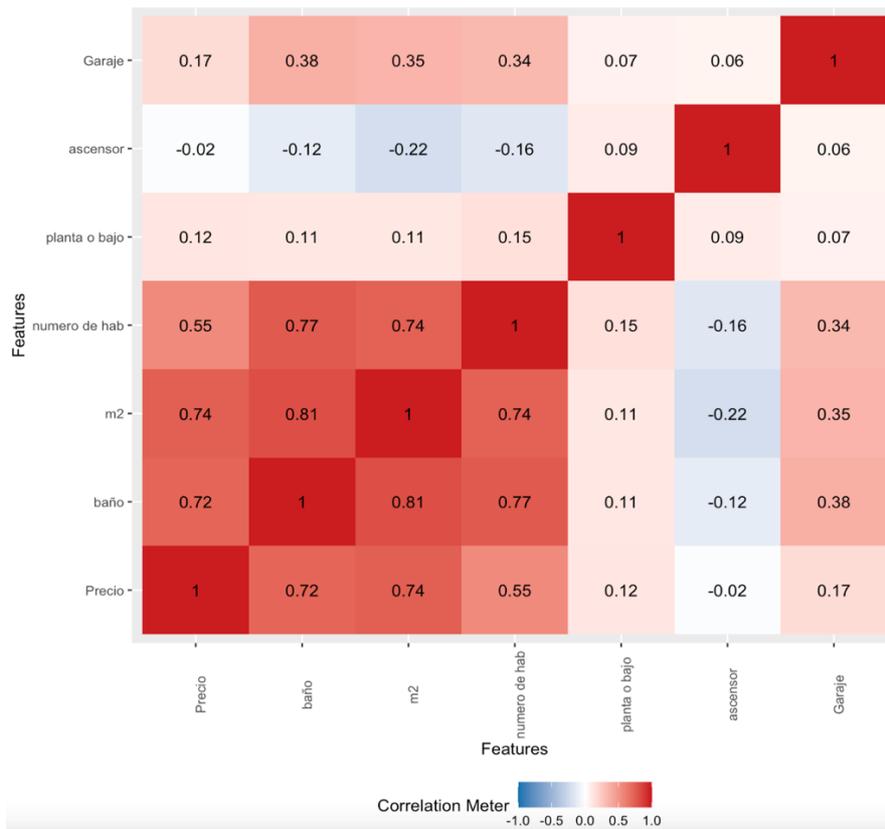
En este apartado se presenta una gráfica con la matriz de correlaciones para conocer el grado de asociación entre las variables. Puede tomar valores desde -1 hasta 1, siendo 1 y -1 el grado de relación más fuerte y 0 el mínimo. Si la correlación es positiva entre dos variables, cuando aumenta o disminuye una la otra también, en el caso de que sea negativa cuando una aumenta la otra disminuye y al revés (Salazar, 2018). Para ello se utilizó la función de correlation en Rstudio En la siguiente tabla se puede apreciar los valores exactos de la correlación entre cada variable y debajo un mapa de calor para entender mejor las relaciones entre variables.

```
#Matriz de correlación para ver si hay correlación entre dos variables
plot_correlation(na.omit(datos), maxcat = 5L)

#tabla de correlaciones
round(cor(x = datos, method = "pearson"), 3)
```

	Precio	baño	m2	num de hab	planta o bajo	ascensor	Garaje
Precio	1.000	0.716	0.744	0.551	0.120	-0.021	0.166
baño	0.716	1.000	0.808	0.769	0.114	-0.124	0.379
m2	0.744	0.808	1.000	0.737	0.109	-0.217	0.349
numero de hab	0.551	0.769	0.737	1.000	0.147	-0.156	0.337
planta o bajo	0.120	0.114	0.109	0.147	1.000	0.089	0.068
ascensor	-0.021	-0.124	-0.217	-0.156	0.089	1.000	0.059
Garaje	0.166	0.379	0.349	0.337	0.068	0.059	1.000

Figura 8: Matriz de correlación entre las variables del dataset



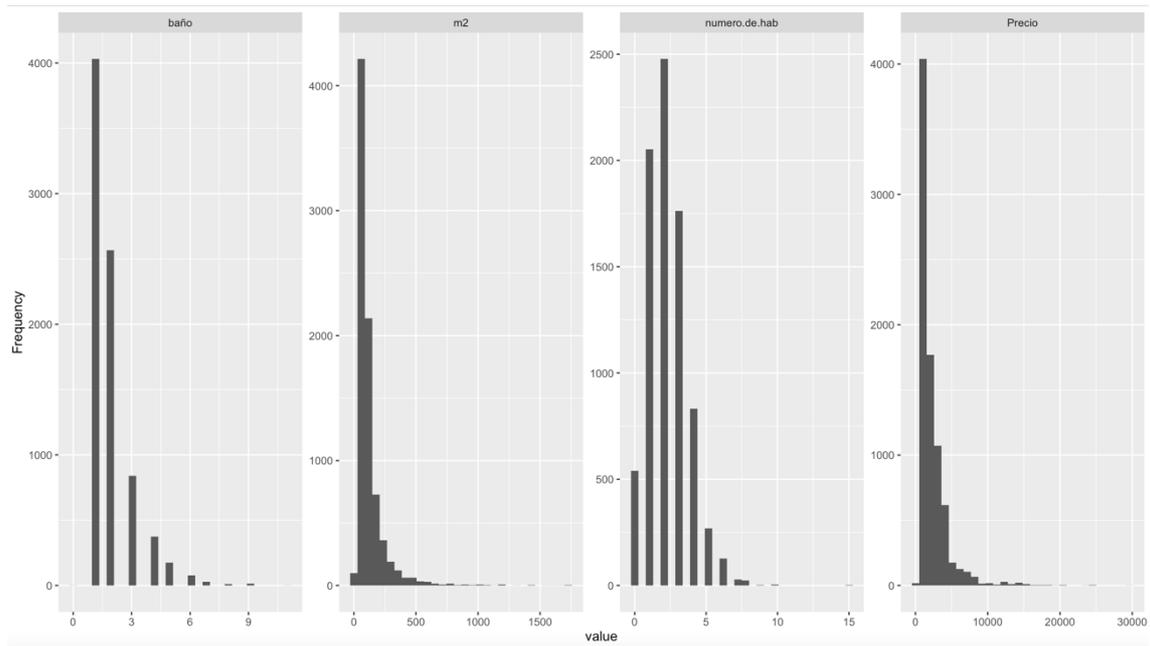
Las variables con una correlación más alta son m² y baño, con un coeficiente de 0,81. Esto significa que cuando una de las dos aumenta la otra también. En cambio, las variables con menor grado de correlación son si tiene ascensor y el precio con valor de 0,02. Podemos observar que las variables que más influenciarán sobre el precio son las que están en rojo, es decir, el número de baños, los m² y el número de habitaciones. Por lo contrario, las que menos influirán son si la vivienda tiene ascensor, si tiene garaje o si el piso se encuentra en un bajo o en una planta.

5.2.3. Distribución de las variables

En este apartado se mostrará un histograma de todas las variables numéricas para ver su distribución. Sólo se representan aquellas variables numéricas que no son dicotómicas, ya que toman únicamente valores de 0 y 1 y no tendrían sentido representarlas.

```
#histograma de todas las variables  
plot_histogram(datos)
```

Figura 9: Histograma de las variables numéricas

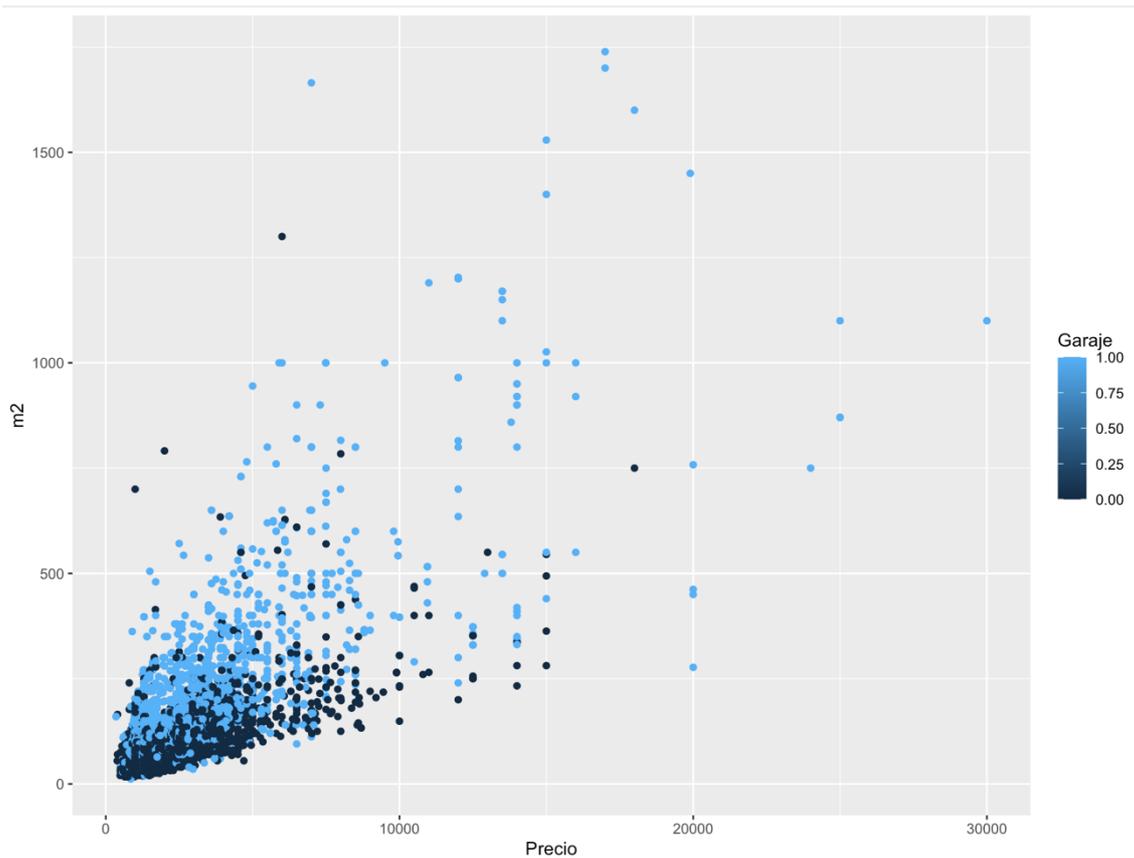


La primera figura representa la distribución de la variable baño, podemos ver que la vivienda tenga entre 1 o 2 baños se repite aproximadamente 4.000 veces, luego la frecuencia va bajando hasta los 6 baños y muy pocos pisos tienen entre 6 y 9 baños. En cuanto a los metros cuadrados alrededor de 4.200 observaciones tienen entre 12 y 100 m² y más de 500 muy pocas. La variable número de habitaciones tiene valores más dispersos, la mayoría de las observaciones tienen entre 1 y 5 habitaciones, la frecuencia de sus observaciones se aproxima a una distribución normal o también llamada distribución gaussiana (Pértegas & Pita, 2001). Si observamos la figura del precio se puede apreciar que la mayoría de las variables están concentradas entre los 350€ y los 2.000€.

A continuación, se muestra un gráfico de dispersión, que sirve para ver la relación entre dos variables numéricas, en el que se observa la relación entre el precio y los metros

cuadrados (Artega et al., 2009). El color indica si esas observaciones tienen garaje o no, siendo azul oscuro aquellos pisos que no tienen garajes y azul claro aquellos que sí.

Figura 10: Gráfico de dispersión entre las variables m2 y precio



Podemos ver que las observaciones con garaje son de aquellos pisos que tienen más metros cuadrados y que las que no están por debajo con menos metros cuadrados. También resaltan las observaciones con precios muy altos y muchos metros cuadrados que todas ellas si tienen garaje.

5.3. Análisis del dataset

En esta parte se analiza el dataset que se ha utilizado para crear el modelo. En primer lugar, utilizando la función que proporciona R Studio para ver si hay valores perdidos, hemos observado que no hay ninguno, por lo que no se ha tenido que hacer ninguna imputación ni tratamiento de N/A.

```
any(is.na(datos))
```

```
FALSE
```

Para ver el tipo de dato de cada variable se ha utilizado la función de R: `str`, que muestra la estructura interna de un objeto.

```
str(datos)
tbl_df [8,118 × 7] (S3: tbl_df/tbl/data.frame)
 $ Precio      : num [1:8118] 950 950 1000 850 1000 870 985 875 975 880 ...
 $ baño        : chr [1:8118] "1" "2" "1" "1" ...
 $ m2          : chr [1:8118] "75" "70" "62" "65" ...
 $ numero de hab: chr [1:8118] "2" "2" "2" "1" ...
 $ planta o bajo: num [1:8118] 1 1 1 1 1 1 1 1 1 1 ...
 $ ascensor    : num [1:8118] 1 1 1 1 1 1 1 1 1 1 ...
 $ Garaje      : num [1:8118] 0 1 0 0 0 0 1 0 1 0 ...
```

Como podemos ver en la tabla las variables baño, m2 y número de habitaciones las coge como categóricas en vez de numéricas, por lo que se procede a cambiar el tipo de variable y a reemplazarlas por las que había.

```
datos$baño=as.numeric(datos$baño)
datos$m2=as.numeric(datos$m2)
datos$`numero de hab`=as.numeric(datos$`numero de hab`)

#comprobamos que todos los datos son numéricos
str(datos)

tbl_df [8,118 × 7] (S3: tbl_df/tbl/data.frame)
 $ Precio      : num [1:8118] 950 950 1000 850 1000 870 985 875 975 880 ...
 $ baño        : num [1:8118] 1 2 1 1 1 1 2 1 1 1 ...
 $ m2          : num [1:8118] 75 70 62 65 42 75 105 61 79 67 ...
 $ numero de hab: num [1:8118] 2 2 2 1 1 2 2 1 2 2 ...
 $ planta o bajo: num [1:8118] 1 1 1 1 1 1 1 1 1 1 ...
 $ ascensor    : num [1:8118] 1 1 1 1 1 1 1 1 1 1 ...
 $ Garaje      : num [1:8118] 0 1 0 0 0 0 1 0 1 0 ...
```

Ahora vemos que las siete variables son de carácter numérico, por lo que son con las que se trabajarán para la creación del modelo.

6. Modelo explicativo

Al tratarse de un problema de predicción de los precios de las viviendas en alquiler de Madrid, se utiliza el algoritmo de regresión lineal múltiple utilizando la técnica de aprendizaje supervisado. En este apartado se va a hacer una regresión como modelo explicativo en el que se va a analizar las variables para ver cuánto impacto tiene un cambio en la variable independiente sobre la dependiente, se usarán todas las observaciones de la base de datos para estimar el modelo y se verá el valor de las betas

6.1. Transformación de variables

En primer lugar, ejecutamos el algoritmo de regresión lineal múltiple con todas las variables sin hacer ninguna modificación en ellas.

```
modelo1<-lm(Precio ~ baño + m2 + `numero de hab` + `planta o bajo` +  
ascensor + Garaje, data = TrainingSet)  
summary(modelo1)
```

	Estimate.	Std. Error	t value	Pr(> t)
(Intercept)	-253.1185	60.6575	-4.173.	3.05e-05 ***
baño	730.6004	24.7912	29.470	< 2e-16 ***
m2	9.3070	0.2174	42.813	< 2e-16 ***
`numero de hab`	-167.3549	18.2419	-9.174	< 2e-16 ***
`planta o bajo`	176.8600	54.1211	3.268	0.00109 **
ascensor	687.9227	37.5131	18.338	< 2e-16 ***
Garaje	-672.7268	34.9320	-19.258	< 2e-16 ***

En la columna de estimate podemos ver para cada variable la beta estimada. Al observar el p-valor, que es la probabilidad de obtener esos datos si la beta realmente fuese 0, podemos evaluar la importancia de esa variable, ya que si es menor que 0.01 será significativo y si es mayor no lo será (Suárez, 2012). Vemos que todas las variables tienen un p-valor >0.01 por lo que todas las variables son importantes para determinar el precio de la vivienda.

Como hemos visto anteriormente en los histogramas de la figura 9 las variables m^2 y precio tienen un gran rango de variación y presentan una asimetría a la derecha. Es por ello que es necesario incluir logaritmos en ambas variables.

```
modelo<-lm(log(Precio) ~ baño + log(m2) + `numero de hab` + `planta o bajo` +
ascensor + Garaje, data = TrainingSet)
summary(modelo)
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	4.426324	0.057374	77.148	<2e-16 ***
baño	0.234112	0.007550	31.009	<2e-16 ***
log(m2)	0.582098	0.015820	36.794	<2e-16 ***
`numero de hab`	-0.080378	0.006342	-12.674	<2e-16 ***
`planta o bajo`	0.142474	0.016496	8.637	<2e-16 ***
ascensor	0.211320	0.011268	18.754	<2e-16 ***
Garaje	-0.284938	0.010799	-26.385	<2e-16 ***

Como podemos ver todas las variables son significativas y el p-valor es menor que en el anterior, por lo que si debemos incluir los logaritmos a ambas variables en el modelo.

6.2. Partición

Una vez transformados los datos para realizar un modelo predictivo es necesario hacer una partición del data set en dos partes: training set y test set. En el training set es donde están los datos que se utilizan para entrenar al modelo y que aprenda los datos, aquí es donde se encuentran la mayor parte de los datos. En cambio, en el test set se encuentran los datos que se utilizan para realizar las predicciones y poder comprobar la performance predictiva del modelo. En este caso se han utilizado el 80% de los datos para la parte de training set y el 20% restante para la parte del test set (IBM, 2023c).

```
TrainIndex<-sample(1:n, n*0.8) #seleccionamos los números de fila de las
observaciones que iran al training set
TrainingSet<-datos[TrainIndex,]
TestSet<-datos[-TrainIndex,]
```

7. Modelo algorítmico predictivo

En este apartado se desarrollan los modelos que se han utilizado para realizar la predicción de los pisos de alquiler de Madrid.

7.1. Regresión lineal múltiple

El primer algoritmo que se ha realizado es una regresión lineal múltiple, que consiste en sustituir en el modelo estimado los valores de sus variables independientes y obtener el valor de la variable a predecir correspondiente. El modelo es lineal ya que son términos aditivos y cada término se multiplica por un coeficiente estimado (IBM, 2023b). En este modelo las betas estimadas son efectos marginales y son constantes, la estimación de las betas se hace a través de mínimos cuadrados. En los modelos lineales la derivada parcial de la variable dependiente, con respecto a cada una de las variables explicativas, es la constante β , y se interpreta como el cambio producido en la variable dependiente cuando la variable independiente aumenta una unidad.

Para realizar el modelo lo primero que se hace es estimar el modelo de regresión sobre el training set con todas las variables seleccionadas tomando logaritmos por la gran diferencia de la magnitud de las variables.

```
modelo<-lm(log(Precio) ~ baño + log(m2) + `numero de hab` + `planta o bajo` +  
ascensor + Garaje, data = TrainingSet)
```

```
summary(modelo)
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	4.426324	0.057374	77.148	<2e-16 ***
baño	0.234112	0.007550	31.009	<2e-16 ***
log(m2)	0.582098	0.015820	36.794	<2e-16 ***
`numero de hab`	-0.080378	0.006342.	-12.674	<2e-16 ***
`planta o bajo`	0.142474	0.016496	8.637	<2e-16 ***
ascensor	0.211320	0.011268	18.754	<2e-16 ***
Garaje	-0.284938	0.010799	-26.385	<2e-16 ***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3771 on 6487 degrees of freedom

Multiple R-squared: 0.6434, Adjusted R-squared: 0.6431

F-statistic: 1951 on 6 and 6487 DF, p-value: < 2.2e-16

El modelo con todas las variables introducidas como predictores tiene un R² considerable de 0.6434, por lo que es capaz de explicar el 64,34% de la variabilidad observada en los precios. El p-valor del modelo es altamente significativo de 2.2e-16, por lo que se puede aceptar que el modelo no es por azar. Todos los p-valor son significativos, lo que es un indicativo de que tienen un impacto significativo en la predicción del precio de las viviendas de alquiler de Madrid.

Para prevenir sobreajuste hay medidas in-sample, basadas en el trainig set que determinan que variables habría que utilizar de las disponibles que hay en el dataset para tener el modelo óptimo. A continuación, se ha hecho otro modelo a través de la metodología stepwise. Es una de las metodologías más simples que consiste en agregar y eliminar de manera iterativa las variables en el modelo para encontrar el subconjunto de variables que mejor ajusta en el conjunto de datos, reduciendo el error de predicción. Se utiliza el número AIC para determinar la calidad del modelo, cuanto menor es el número, mejor es el modelo, ya que indica que el modelo proporciona una buena explicación de los datos con un número mínimo de parámetros. La estrategia stepwise comienza sin predictores y se van añadiendo los predictores más contributivos. A cada vez que se agrega una variable, se elimina otra que no proporcione mejora en el ajuste del modelo (Mendez, 2019).

El segundo modelo se ha creado utilizando la estrategia stepwise.

```
stepmodel<-step(object= modelo, direction="both", trace=1)
```

```
summary(stepmodel)
```

```
# obtenemos predicciones en el TestSet
```

```
predicciones<-predict(modelo, newdata = TestSet)
```

La combinación de variables utilizando la estrategia stepwise y con criterio matemático de menor AIC es la misma que la anterior, es decir, utilizando todas las variables. Como hemos visto anteriormente todas son muy significativas y contribuyen al cálculo de la variable dependiente.

```
Start: AIC=-12658.94
```

```
log(Precio) ~ baño + log(m2) + `numero de hab` + `planta o bajo` +  
ascensor + Garaje
```

El valor de AIC y BIC coinciden para ambos modelos, ya que son los mismo y utilizan todas las variables.

```
> AIC(fullmodel)
```

```
[1] 5772.233
```

```
> AIC(stepmodel)
```

```
[1] 5772.233
```

```
> BIC(fullmodel)
```

```
[1] 5826.462
```

```
> BIC(stepmodel)
```

```
[1] 5826.462
```

Una vez terminado el modelo se va a evaluar la performance predictiva, para ello se va a medir el error de predicción que es la diferencia entre el valor real y el valor predicho por el modelo. Esto se hace con los datos del test set que no han sido usado para la creación del modelo. En primer lugar, se obtienen las predicciones del test set y después se calcula el error entre las predicciones del test set y los valores reales del test set de la siguiente manera (MathWorks, n.d.).

```
# obtenemos predicciones en el TestSet
```

```
predicciones<-predict(modelo, newdata = TestSet)
```

```
# calculamos los errores de predicción
```

```
errores<-TestSet$Precio-predicciones
```

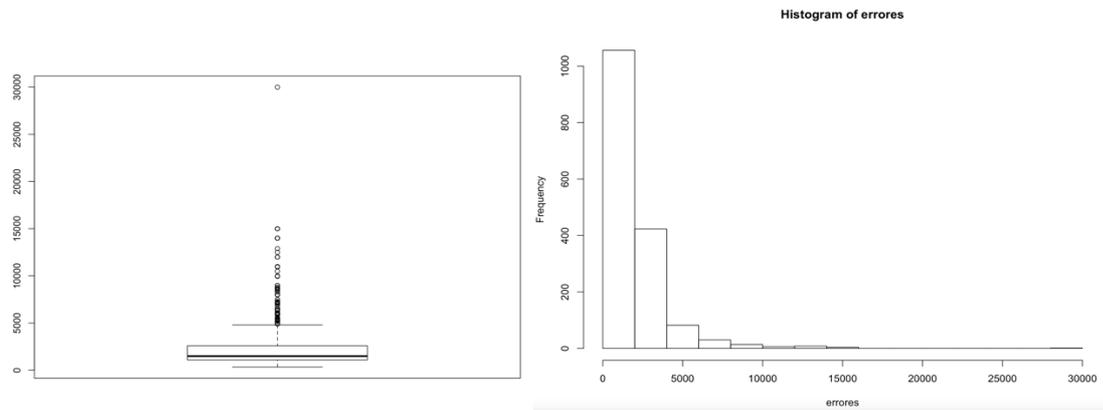
```
boxplot(errores) # gráfica de los errores de predicción
```

```
hist(errores, breaks=20)
```

```
h<-nrow(TestSet) # numero de observaciones en el TestSet
```

En las siguientes figuras podemos ver que los errores de predicción del modelo se encuentran alrededor de los 1000, que es un valor bastante alto. También podemos identificar un outlier con un error de predicción de 3000.

Figura 11: Histograma y boxplot de los errores de predicción



Por último, se calculan una serie de medidas de performance predictiva que son el RMSE que es la raíz del error cuadrático medio, el MAPE que es el error absoluto porcentual medio y el MAE que es el error absoluto medio. Cuanto menor sean estos errores mejor es el modelo, el objetivo es conseguir un algoritmo con el menor error (IBM, 2023a).

```
# cálculo de las medidas de performance predictiva en el TestSet
ME<-sum(errores)/h
RMSE<-sqrt(sum(errores^2)/h)
MAE<-sum(abs(errores))/h
MAPE<-sum (abs(errores)/TestSet$Precio)/h*100
> ME
[1] 2176.818
> RMSE
[1] 2936.262
> MAE
[1] 2176.818
> MAPE
[1]99.51287
```

7.2. Random forest

El segundo algoritmo que se ha realizado es un random forest, que en realidad es un ensemble de varios árboles de decisión mediante la técnica de bagging. En este algoritmo lo que se hace es que se extraen muchas muestras aleatorias del training set con reemplazamiento de datos. Así se reduce la varianza del modelo, ya que se entrena con muchos más datos y hace que el modelo generalice mejor. Para ello se utiliza un subconjunto aleatorio de predictores por cada una de las divisiones y se obtiene un árbol para cada una de las muestras aleatorias, es por ello que se acaba obteniendo un conjunto de árboles de decisión, denominado bosque. Se combinan las predicciones de todos los árboles haciendo un promedio de todas ellas (IBM, n.d.).

Para la creación y entrenamiento del modelo se utiliza la función de Rstudio, ranger, poniendo la variable dependiente en la fórmula y utilizando los datos del training set. Al igual que en el algoritmo anterior se utilizan los datos de training set para entrenar al modelo y los de test set para hacer las predicciones y medir la performance predictiva.

```
modelo <- ranger(  
  formula = Precio ~ .,  
  data = TrainingSet,  
  num.trees = 10,  
  seed = 123)  
print(modelo)
```

Ranger result

Call:

```
ranger(formula = Precio ~ ., data = TrainingSet, num.trees = 10, seed = 123)
```

Type: Regression

Number of trees: 10

Sample size: 6494

Number of independent variables: 6

Mtry: 2

Target node size: 5

Variable importance mode: none

```
Splitrule:          variance
OOB prediction error (MSE):  1383157
R squared (OOB):          0.6717205
```

Se han incluido 10 árboles en el modelo, como dataset se ha cogido las observaciones del training set que son 6.494, las 6 variables independientes del modelo. Se han considerado 2 predictores por cada división, que es la raíz cuadrada del número de predictores disponibles redondeado hacia abajo.

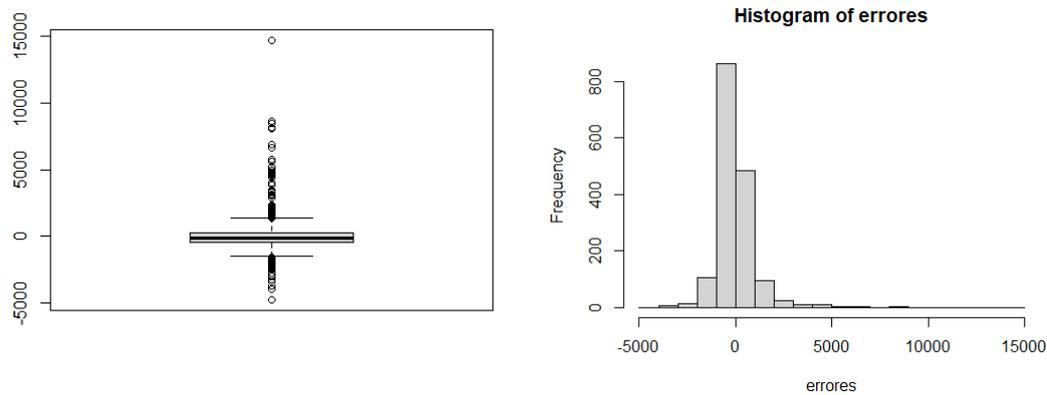
Por último, se ha evaluado la capacidad predictiva del modelo con los datos del test set. Para ello se han calculado las predicciones de las observaciones del test set y después se ha calculado el error restando el valor real de las observaciones del test set menos el valore predicho por el modelo (IBM, 2023a).

```
predicciones <- predict(modelo,data = TestSet)
predicciones <- predicciones$predictions
test_rmse <- sqrt(mean((predicciones - TestSet$Precio)^2))

# calculamos los errores de predicción
errores<-TestSet$Precio-predicciones
boxplot(errores) # gráfica de los errores de predicción
hist(errores, breaks=20)
h<-nrow(TestSet) # numero de observaciones en el TestSet
```

En las siguientes gráficas podemos ver que los errores de predicción del modelo se encuentran alrededor de 0, por lo que es bastante bueno. También podemos identificar que hay algunos outliers.

Figura 12: Histograma y boxplot de los errores de predicción



```
# cálculo de las medidas de performance predictiva en el TestSet
```

```
ME<-sum(errores)/h
```

```
RMSE<-sqrt(sum(errores^2)/h)
```

```
MAE<-sum(abs(errores))/h
```

```
MAPE<-sum (abs(errores)/TestSet$Precio)/h*100
```

```
> ME
```

```
[1] -1.611
```

```
> RMSE
```

```
[1] 1088.705
```

```
> MAE
```

```
[1] 637.908
```

```
> MAPE
```

```
[1] 31.884
```

8. Resumen e interpretación de los resultados

En este epígrafe se hará un resumen de los resultados obtenidos en el apartado anterior para los dos algoritmos y se interpretarán los resultados de ambos. Es importante medir la performance predictiva de cada algoritmo para ver si son precisos y poder elegir el que mejor estime los datos entre los dos.

Los resultados de las medidas de performance predictiva del test set para el primer algoritmo que es el de regresión lineal múltiple, fueron los siguientes:

```
> ME
[1] 2176.818
> RMSE
[1] 2936.262
> MAE
[1] 2176.818
> MAPE
[1]99.51287
```

Con esto podemos deducir que las predicciones del modelo tienen un error medio cuadrático de 2939,26 unidades del precio de alquiler. Si vemos el MAPE (error medio absoluto porcentual) podemos deducir que las predicciones tienden a desviarse un 99,51% de los valores reales. A su vez, si nos fijamos en el error absoluto y estándar de la media, vemos que es muy alto, con un valor de 2176,68. Estos resultados revelan que las predicciones tienen un margen de error considerable.

El segundo modelo fue entrenado con un random forest y a continuación veremos los resultados de la evaluación de la performance predictiva, para así poder compararlos con el modelo anterior y elegir el que mejor se ajuste.

```
> ME
[1] -1.611
> RMSE
[1] 1088.705
> MAE
[1] 637.908
> MAPE
[1] 31.884
```

Dados estos datos podemos concluir que las predicciones tienen un error medio cuadrático de 1088,70 unidades del precio de alquiler. En cuanto al error absoluto porcentual medio es de 31,88%, por lo que el 32% de las observaciones tienden a desviarse de sus valores reales. El error medio tiene un valor bajo y a la vez negativo, esto significa que las predicciones suelen ser menores que los valores reales, pero sólo en

1,61 unidades. Estos resultados indican que el modelo proporciona predicciones bastante precisas.

En conclusión, observando los resultados de ambos algoritmos es mejor el modelo que se entrena con el algoritmo de random forest que el de regresión lineal múltiple, ya que todas las medidas de performance predictiva del algoritmo random forest presentan valores menores que el de regresión, por lo que los valores predichos se acercan más a los valores reales de las observaciones. Esto se debe a que el algoritmo de random forest es menos sensible a valores atípicos y es capaz de captar mejor las relaciones entre las variables (IBM, n.d.).

9. Conclusión

Esta tesis aborda la complejidad y el cambio que ha experimentado el mercado inmobiliario español durante las últimas décadas. El mercado ha demostrado su capacidad para adaptarse a desafíos, incluida la pandemia de COVID-19, desde la época de gran prosperidad y crecimiento económico hasta la crisis económica de 2008 y la actualidad. La planificación urbana sostenible, la transparencia en las transacciones inmobiliarias y la consideración de factores globales se demuestran cruciales para garantizar un futuro crecimiento justo y sostenible. El éxito del mercado inmobiliario español dependerá también de la colaboración entre los sectores público y privado, así como de la capacidad de adaptarse a las tendencias globales.

El histórico de los precios de las viviendas de alquiler en España muestra una gran fluctuación a lo largo de los años. Esta variabilidad ha sido causada por la oferta y la demanda, las políticas gubernamentales, los cambios en los tipos de interés, la inversión extranjera y los ciclos económicos. Los precios actualmente alcanzan máximos históricos, y se prevé una corrección en 2024. El análisis de factores macroeconómicos revela la compleja interacción entre la inflación, los costos de construcción y los tipos de interés. El conflicto ruso-ucraniano y la pandemia elevaron los precios de las materias primas y la cadena de suministro, lo que ha resultado en el aumento de los costes de construcción. Se espera que el equilibrio de precios se alcance alrededor de 2025, lo que impulsará la inversión extranjera en la construcción de viviendas. Los tipos de interés también juegan un papel importante, ya que afectan la demanda de propiedades y la

accesibilidad financiera. La reciente ley de vivienda de 2023 introduce cambios significativos que afectan tanto a los inquilinos como a los propietarios en términos de precios y beneficios.

Para comprender las tendencias del mercado se ha desarrollado un modelo predictivo de precios de alquiler de viviendas en Madrid mediante algoritmos. El uso de un modelo predictivo y datos históricos permite anticipar y comprender posibles fluctuaciones en los precios de las viviendas. Este método facilita la toma de decisiones informadas para los participantes del mercado inmobiliario.

La exploración de datos y el modelado algorítmico muestran un análisis detallado de múltiples factores que afectan a los precios de alquiler de viviendas en Madrid. Algoritmos como random forest y la regresión lineal múltiple han permitido predecir los precios de las viviendas en alquiler de Madrid. La inclusión de técnicas de selección de variables y la exploración de transformaciones logarítmicas buscan mejorar el modelo.

El análisis descriptivo presentado enfatiza que variables como el número de baños, los metros cuadrados y el número de habitaciones tienen un alto impacto sobre el precio. Además, el uso del algoritmo random forest ofrece una perspectiva adicional, demostrando la utilidad de métodos más complejos para predecir precios.

Sin embargo, los hallazgos muestran que hay limitaciones y áreas de mejora en cuanto al modelo predictivo. El análisis de errores nos hace pensar que cabría la posibilidad de mejorar los modelos y resolver sesgos potenciales.

Estos resultados preparan las bases para investigaciones y mejoras futuras, enfatizando la importancia de comprender profundamente las complejidades del mercado inmobiliario para crear modelos predictivos sólidos y aplicables.

En conclusión, el estudio expuesto ofrece una perspectiva completa del mercado inmobiliario español, analizando su desarrollo histórico y las dificultades actuales. Es importante tener estrategias de planificación urbana sostenible, transparencia en las transacciones inmobiliarias y adaptabilidad a los factores globales para asegurar un crecimiento justo y sostenible. El análisis detallado de la fluctuación de los precios de

alquiler muestra la compleja interacción entre los factores macroeconómicos y las tendencias locales. No obstante, se reconoce la importancia de abordar las limitaciones identificadas, el modelo predictivo desarrollado ofrece una herramienta valiosa para comprender y anticipar las dinámicas del mercado en Madrid. Finalmente, esta investigación ayuda a comprender el mercado inmobiliario español y establece una base sólida para futuras investigaciones y mejoras en la predicción de precios de vivienda.

10. Bibliografía

- Alves, P., & San Juan, L. (2021). El impacto de la crisis sanitaria del COVID-19 sobre el mercado de la vivienda en España. *Boletín Económico*, (2/2021).
- ACR. (2022). (rep.). *Índice de Costes Directos de Construcción ACR*. Madrid
- Arellano, M., & Bentolila, S. (2009). La burbuja inmobiliaria: causas y responsables. *La Crisis de la Economía Española: Lecciones y Propuestas*. FEDEA, <http://www.crisis09.es/ebook>.
- Arteaga, P., Batanero, C., Díaz, C., & Contreras, J. M. (2009). El lenguaje de los gráficos estadísticos. *UNION-Revista Iberoamericana de Educación Matemática*, 5(18).
- Bernardos Domínguez, G. (2009). Creación y destrucción de la burbuja inmobiliaria en España. *Información Comercial Española. Revista de Economía ICE*, 2009, num. 850, p. 23-40.
- BOE. (2019). *Agencia Estatal Boletín Oficial del Estado*. Ir a la página de inicio. <https://www.boe.es/buscar/act.php?id=BOE-A-2019-3814>
- BOE. (2023). *Agencia Estatal Boletín Oficial del Estado*. Legislación consolidada. <https://www.boe.es/buscar/act.php?id=BOE-A-2023-12203>
- Banegas, J. S. (2022, December 23). *¿Cómo evolucionarán los costes de la construcción en España en 2023?*. CaixaBank Research. <https://www.caixabankresearch.com/es/analisis-sectorial/inmobiliario/como-evolucionaran-costes-construccion-espana-2023>
- CBRE. (2022, September 6). Costes de Construcción: Hacia la senda de la recuperación. <https://www.cbre.es/insights/articles/costes-de-construccion-hacia-la-senda-de-la-recuperacion>

CBRE. (2022). *Impacto Del Aumento de Los Costes de Construcción En El Sector*
www.cbre.es/insights/reports/impacto-del-aumento-de-los-costes-de-construccion-en-el-sector-inmobiliario.

CTE. (2019). DB Se Diciembre2019 - codigotecnico.org.
<https://www.codigotecnico.org/pdf/Documentos/SE/DBSE.pdf>

de la Paz, P. T. (2006). La oferta de viviendas y el mercado inmobiliario en España.
Papeles de Economía Española, 109, 156-181.

Harris, J. C. (1989). The effect of real rates of interest on housing prices. *The Journal of Real Estate Finance and Economics*, 2, 47-60.

IBM. (2023a). *Forecasting statistical details*. <https://www.ibm.com/docs/es/cognos-analytics/11.1.0?topic=forecasting-statistical-details>

IBM. (2023b). *Multiple linear regression*. <https://www.ibm.com/docs/es/cognos-analytics/11.1.0?topic=tests-multiple-linear-regression>

IBM. (2023c). Regresión Lineal ridge. <https://www.ibm.com/docs/es/spss-statistics/saas?topic=features-linear-ridge-regression>

IBM. (n.d.). ¿Qué es el aprendizaje supervisado? <https://www.ibm.com/es-es/topics/supervised-learning>

IBM. (n.d.). ¿Qué es un bosque aleatorio? <https://www.ibm.com/es-es/topics/random-forest>

Idealista. (2007). Informe de precios en alquiler en España — IDEALISTA.
<https://www.idealista.com/sala-de-prensa/informes-precio-vivienda/alquiler/report/>

Idealista. (2014). Informe de Precios de Alquiler en España de 2014 — IDEALISTA.
<https://www.idealista.com/sala-de-prensa/informes-precio-vivienda/alquiler/report/2014/>

Idealista. (2023, October 2). El Precio del Alquiler Sigue subiendo en las grandes ciudades con la Ley de Vivienda en vigor.

<https://www.idealista.com/news/inmobiliario/vivienda/2023/10/02/808325-la-ley-de-vivienda-no-consigue-frenar-la-escalada-de-precios-del-alquiler-en-las>

Idealista. (2022, March 6). *España Necesita 1,2 millones de viviendas de alquiler para cubrir La Demanda, Según Savills.* idealista/news. <https://www.idealista.com/news/inmobiliario/vivienda/2022/03/04/795270-economia-savills-aguirre-newman-estima-que-se-necesitan-1-2-millones-de-viviendas-mas-para-cubrir-la#:~:text=En%20España%20hay%204%2C2,2030%2C%20según%20Savills%20Aguirre%20Newman>

Idealista. (2023a). Evolución del Precio de la Vivienda en alquiler en Madrid - IDEALISTA. <https://www.idealista.com/sala-de-prensa/informes-precio-vivienda/alquiler/madrid-comunidad/madrid-provincia/madrid/>

Idealista. (2023b). Evolución del Precio de la Vivienda en alquiler en Toledo - IDEALISTA. <https://www.idealista.com/sala-de-prensa/informes-precio-vivienda/alquiler/castilla-la-mancha/toledo-provincia/toledo/>

Idealista. (2023c). Informe de precios en alquiler en España — IDEALISTA. <https://www.idealista.com/sala-de-prensa/informes-precio-vivienda/alquiler/report/>

Idealista. (2023d, October 10). *Precio de la Vivienda en 2024: Los Expertos Auguran “cautela” y una “Corrección” A Nivel Nacional.* idealista/news. <https://www.idealista.com/news/inmobiliario/vivienda/2023/10/10/808488-precio-de-la-vivienda-en-2024-los-expertos-auguran-cautela-y-una-correccion-a-nivel>

Knight Frank. (2021). Global Residential Cities Index 2021.

Llinás, H. (2021). Estadística descriptiva en R.

https://rstudio-pubs-static.s3.amazonaws.com/647537_aa2563a0272f4e7dbc8852012951a00e.html

- López Rodríguez, D., & Matea Rosa, M. D. L. L. (2019). Evolución reciente del mercado del alquiler de vivienda en España. *Boletín económico/Banco de España*, n. 3, 2019.
- Martinez, J., & Maza, L. A. (2003). Análisis del Precio de la Vivienda en España. <https://repositorio.bde.es/bitstream/123456789/6770/1/dt0307.pdf>
- MathWorks. (n.d.). *Regresión lineal*. Regresión Lineal - Matlab & Simulink - Mathworks España. https://es.mathworks.com/help/matlab/data_analysis/linear-regression.html
- Mayordomo Tello, P. (2019). Causas Que Afectan a la Subida de Precios de Alquiler y Alternativa Entre Comprar o Alquilar Vivienda en la Actualidad
- Mendez, J. (2019). *Stepwise Regresión*. RPubS. https://rpubs.com/jorge_mendez/609253
- Montalvo, J. G. (2007). El sector inmobiliario español a principios del siglo XXI: entre la demografía y las expectativas. *Cim. economía: Revista económica de Castilla-La Mancha*, 11, 57-79.
- Montoriol, J. (2022, December 23). El Sector Inmobiliario Se Enfria. <https://www.caixabankresearch.com/es/analisis-sectorial/inmobiliario/sector-inmobiliario-se-enfria>
- Pértegas Díaz, S., & Pita Fernández, S. (2001). La distribución normal. *Cad Aten Primaria*, 8, 268-274.
- Polanco, C. (2023, May 4). *El Precio de la Vivienda, en máximos históricos en cuatro capitales españolas*. EXPANSION. <https://www.expansion.com/economia/2023/05/04/64536a22468aeb412c8b4645.html>
- Reichert, A. K. (1990). The impact of interest rates, income, and employment upon regional housing prices. *The Journal of Real Estate Finance and Economics*, 3, 373-391.

Salazar, C. (2018). *Cálculo de correlaciones*. RPubS.
<https://rpubs.com/camilamila/correlaciones>

Savills. (2022, March 7). Savills Aguirre Newman estima que son necesarias 1,2 millones de viviendas adicionales en Alquiler. Ir a inicio.
<https://www.savills.es/publicaciones/noticias/325875-0/savills-estima-que-son-necesarias-1-2-millones-de-viviendas-adicionales-en-alquiler>

Suárez, N. R. (2012). La revolución en la toma de decisiones estadísticas: el p-valor. *Telos*, 14(3), 439-446.

Trecet, J. (2021, June 11). *Así es el inquilino perfecto por quienes Alquilan de Forma Profesional*. Business Insider España. <https://www.businessinsider.es/inquilino-perfecto-quienes-alquilan-forma-profesional-545813>

Wan, L. J., Zhang, Y., & Wang, G. X. (2009, September). Macro-micro factors affecting real estate demand analysis. In 2009 International Conference on Management Science and Engineering (pp. 1988-1993). IEEE.