



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA

ADE

Facultad de Administración  
y Dirección de Empresas /UPV

UNIVERSITAT POLITÈCNICA DE VALÈNCIA

Facultad de Administración y Dirección de Empresas

Determinación de la rentabilidad en el mercado de naves  
industriales de la Comunidad Valenciana en 2022

Trabajo Fin de Grado

Grado en Administración y Dirección de Empresas

AUTOR/A: Mrozowski Handzel, Arturo

Tutor/a: Ribal Sanchis, Francisco Javier

CURSO ACADÉMICO: 2022/2023

## Resumen

---

La ratio de rentabilidad bruta del alquiler es un indicador utilizado de forma habitual por los inversores para analizar la rentabilidad de bienes inmuebles. En este sentido, la proliferación y popularidad de portales de anuncios inmobiliarios en Internet permite obtener información suficiente para que el cálculo de la rentabilidad bruta y su posterior estudio tenga relevancia. El objetivo del presente Trabajo Final de Grado (TFG) es determinar la rentabilidad bruta del alquiler de naves industriales en la Comunidad Valenciana. Para ello, se parte de una base de datos de ofertas de venta y alquiler disponibles en un portal inmobiliario construida mediante técnicas de *web scraping*, extrayéndose de los anuncios variables como el precio en euros, la superficie y la localización, entre otras. La base de datos se depura eliminando aquellas ofertas anómalas que podrían afectar a la obtención de resultados estadísticamente significativos y posteriormente se aplican técnicas estadísticas para analizar la oferta de naves en función de sus características y descubrir aquellos patrones que podrían condicionar la estimación de la distribución de la ratio de rentabilidad a partir de la muestra.

## Resum

---

La ràtio de rendibilitat bruta del lloguer és un indicador utilitzat habitualment pels inversors per analitzar la rendibilitat de béns immobles. En aquest sentit, la proliferació i popularitat de portals d'anuncis immobiliaris a Internet permet obtenir prou informació perquè el càlcul de la rendibilitat bruta i el seu posterior estudi tinga rellevància. L'objectiu del present Treball Final de Grau (TFG) és determinar la rendibilitat bruta del lloguer de naus industrials a la Comunitat Valenciana. Per això, es construïx mitjançant tècniques de *web scraping* una base de dades d'ofertes de venda i lloguer disponibles en un portal immobiliari, extraient-se dels anuncis variables com el preu en euros, la superfície i la localització, entre d'altres. La base de dades es depura eliminant aquelles ofertes anòmales que podrien afectar l'obtenció de resultats estadísticament significatius i posteriorment s'apliquen tècniques estadístiques per a analitzar l'oferta de naus en funció de les seues característiques i descobrir aquells patrons que podrien condicionar l'estimació de la distribució de la ràtio de rendibilitat a partir de la mostra.

## Abstract

---

The rent-to-price (RTP) ratio is an indicator commonly used by investors to analyze the return of real estate investments. In this sense, the proliferation and popularity of real estate listings websites on the Internet allows to obtain sufficient information, so the calculation of the gross return and its subsequent study has relevance. The objective of this Bachelor's Thesis is to determine the RTP ratio of industrial buildings in the Valencian Community (Spain). To do so, a database of sale and rent offers available in a real estate listings website is built by web scraping techniques, extracting from the listings variables such as price in euros, area, and location, among others. The database is cleaned by eliminating those anomalous offers that could affect obtaining statistically significant results and then statistical techniques are applied to analyze the industrial buildings market according to their characteristics and to discover those patterns that could condition the estimation of the distribution of the RTP ratio from the sample.

# Índice general

---

<b>1</b>	<b>Introducción</b>	<b>1</b>
1.1	Contexto . . . . .	1
1.2	Motivación . . . . .	5
1.3	Objetivos . . . . .	6
1.4	Orden documental . . . . .	7
<b>2</b>	<b>Marco conceptual</b>	<b>9</b>
2.1	Rentabilidad en inversiones inmobiliarias . . . . .	9
2.1.1	Concepto de inversión . . . . .	9
2.1.2	Concepto de rentabilidad . . . . .	11
2.1.3	Ratio de rentabilidad bruta del alquiler . . . . .	12
2.2	<i>Web scraping</i> . . . . .	13
2.2.1	Concepto de <i>web scraping</i> . . . . .	13
2.2.2	Procedimiento de extracción de datos . . . . .	14
2.2.3	Ética y legalidad del <i>web scraping</i> . . . . .	15
<b>3</b>	<b>Metodología</b>	<b>17</b>
3.1	Obtención de información . . . . .	18
3.1.1	Elección de la fuente de información . . . . .	18
3.1.2	Implementación del <i>web scraper</i> . . . . .	21
3.2	Construcción de la base de datos . . . . .	23
3.2.1	Generación de nuevas variables . . . . .	24
3.2.2	Limpieza de datos . . . . .	26
3.2.3	Estructura final de la base de datos . . . . .	30
3.3	Análisis exploratorio . . . . .	31
3.3.1	Distribución geográfica . . . . .	31
3.3.2	Precio . . . . .	36
3.3.3	Superficie . . . . .	40
3.4	Determinación de la rentabilidad bruta . . . . .	46
<b>4</b>	<b>Resultados y discusión</b>	<b>47</b>
4.1	Rentabilidad bruta del alquiler por provincia . . . . .	47
4.2	Rentabilidad bruta del alquiler por comarca . . . . .	48
4.3	Rentabilidad bruta del alquiler en las capitales de provincia . . . . .	52
<b>5</b>	<b>Conclusiones y propuestas de trabajo futuro</b>	<b>53</b>
	<b>Bibliografía</b>	<b>57</b>

## **Anexos**

### **A Objetivos de Desarrollo Sostenible**

**61**

# Índice de figuras

---

1.1	Ratio entre compraventas anuales de inmuebles comerciales y viviendas en España	2
1.2	Transacciones trimestrales en el sector inmobiliario comercial según el tipo de inmueble . . . . .	3
1.3	Evolución anual del precio medio de locales y naves en España (en €/m <sup>2</sup> ) . . . . .	4
2.1	Etapas del <i>web scraping</i> . . . . .	14
3.1	Metodología de trabajo . . . . .	17
3.2	Página principal de Idealista . . . . .	19
3.3	Mapa de la provincia de Valencia en Idealista . . . . .	19
3.4	Listado de anuncios de naves en venta en Idealista en la provincia de Valencia .	20
3.5	Diagrama de bloques del <i>web scraper</i> implementado para recopilar información de Idealista . . . . .	22
3.6	Nube de palabras con los 100 términos más repetidos en las descripciones de los anuncios . . . . .	25
3.7	Diagrama de caja y bigotes de la variable superficie de las naves en venta . . . . .	28
3.8	Distancia de Mahalanobis aplicada a la muestra de naves a la venta en la costa de Alicante . . . . .	29
3.9	Intensidad de la oferta de naves industriales en las provincias de la Comunidad Valenciana . . . . .	32
3.10	Intensidad de la oferta de naves industriales en la Comunidad Valenciana por comarca . . . . .	34
3.11	Curva de densidad de la variable precio . . . . .	38
3.12	Precio medio de las naves industriales por comarca . . . . .	38
3.13	Curvas de densidad de la variable superficie en el mercado de oferta y de alquiler	41
3.14	Superficie media de las naves industriales por comarca . . . . .	42
3.15	Dispersión de la superficie de las naves industriales en venta y en alquiler por comarca . . . . .	45
4.1	Mapa de rentabilidad media del alquiler de las naves industriales por comarca .	50

# Índice de tablas

---

3.1	Datos recopilados en el proceso de <i>web scraping</i> . . . . .	23
3.2	Resumen de las variables registradas en la base de datos procesada . . . . .	30
3.3	Oferta de naves industriales por provincia en la Comunidad Valenciana . . . . .	33
3.4	Distribución por comarcas de la oferta de naves industriales de la Comunidad Valenciana . . . . .	35
3.5	Oferta de naves industriales en las capitales de provincia de la Comunidad Valenciana	36
3.6	Resumen de la variable precio de las naves en venta . . . . .	37
3.7	Resumen de la variable precio de las naves en alquiler . . . . .	37
3.8	Precio de la oferta de naves industriales por comarca . . . . .	39
3.9	Resumen de la variable superficie de las naves en venta . . . . .	40
3.10	Resumen de la variable superficie de las naves en alquiler . . . . .	40
3.11	Superficie de la oferta de naves industriales por comarca . . . . .	43
4.1	Índices de rentabilidad media por provincia . . . . .	47
4.2	Índices de rentabilidad mediana por provincia . . . . .	47
4.3	Índices de rentabilidad media por comarca . . . . .	48
4.4	Índices de rentabilidad mediana por comarca . . . . .	49
4.5	Índices de rentabilidad media por comarcas de costa e interior . . . . .	51
4.6	Índices de rentabilidad mediana por comarcas de costa e interior . . . . .	51
4.7	Índices de rentabilidad media por capitales de provincia . . . . .	52
4.8	Índices de rentabilidad mediana por capitales de provincia . . . . .	52





## 1.1. Contexto

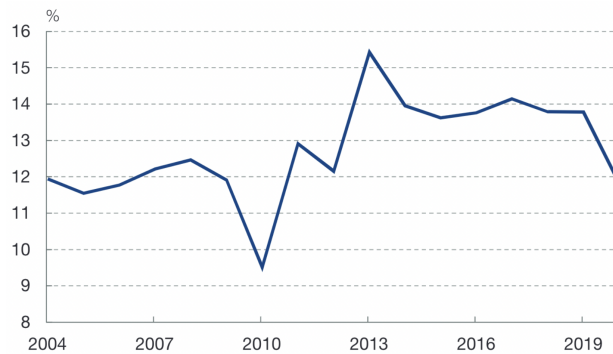
El mercado inmobiliario constituye un sector clave en la economía de un país y, como tal, su análisis es amplio y permanente. Dentro del mercado inmobiliario, las transacciones de viviendas tienen un peso y atención superiores en comparación con las negociaciones sobre cualquier otra categoría de bienes inmuebles (Casas del Rosal, Caridad y Ocerín & Núñez Tabales, 2018). Esta notable importancia del mercado residencial con respecto al resto del sector estriba en que se trata del mayor mercado tanto en términos de oferta, entendida esta como el número de propiedades disponibles, como de demanda, puesto que una vivienda cubre una necesidad básica de las personas. Precisamente, este carácter de bien esencial de la vivienda provoca que este mercado se halle especialmente sujeto a regulación, hecho que también contribuye a su peso superlativo en la economía.

Debido al predominio de las viviendas y la atención que se le presta a dicho sector en los estudios de índole económica, es común identificar el mercado inmobiliario únicamente con el mercado de la vivienda. Sin embargo, al margen del mercado residencial, el mercado inmobiliario también abarca inmuebles en los que se desarrollan actividades empresariales, como oficinas, locales comerciales o naves industriales. De esta forma, se puede establecer una analogía entre el mercado residencial, al que acuden los agentes económicos familias para cubrir su necesidad habitacional, y el mercado de esta tipología de inmuebles de uso terciario o mercado comercial, al que acuden los agentes económicos empresas buscando disponer de un lugar físico en el que desarrollar su actividad económica.

Si bien muchas transacciones inmobiliarias tienen lugar por los motivos expuestos previamente, el mercado inmobiliario puede entenderse, además, desde un punto de vista muy diferente dado que también representa una alternativa de inversión. En efecto, un inversor que diseña su cartera de activos de inversión no está necesariamente limitado a los productos de inversión disponibles en los mercados financieros (activos de renta fija y variable, divisas, etc.), sino que también puede incluir activos inmobiliarios en la misma (Amargant, 2008). La composición de la cartera de inversión viene determinada por el perfil de riesgo asociado al inversor y, en general, la evolución del mercado inmobiliario es más estable y predecible que en muchos mercados financieros, por lo que se vincula a perfiles más conservadores.

Cabe resaltar que el mercado inmobiliario también se caracteriza por requerir una inversión inicial más elevada que algunos activos disponibles en el mercado financiero. No obstante, lo que inicialmente podría parecer una barrera de entrada no ha supuesto un impedimento en los últimos años, dada la política de tipos de interés cero desarrollada por el Banco Central Europeo con el objetivo de aumentar la disponibilidad de crédito para fomentar el consumo y la inversión (Banco Central Europeo, 2022).

Habida cuenta de los activos inmobiliarios como posible elección a la hora de realizar una inversión, es conveniente profundizar en la estructura del sector inmobiliario español y conocer su situación. Tal y como ya se ha apuntado, el predominio del mercado de la vivienda en el conjunto del mercado inmobiliario es igualmente notorio en España. En la [Figura 1.1](#) puede observarse que, entre 2014 y 2019, las transacciones en el mercado inmobiliario comercial suponían alrededor de un 14 % de las operaciones anuales de compraventa de viviendas (Lamas & Romaniega, 2022). Recientemente, destaca el descenso del año 2020 debido fundamentalmente a la crisis sanitaria originada por la COVID-19.



**Figura 1.1:** Ratio entre compraventas anuales de inmuebles comerciales y viviendas en España

**Fuente:** Lamas y Romaniega (2022)

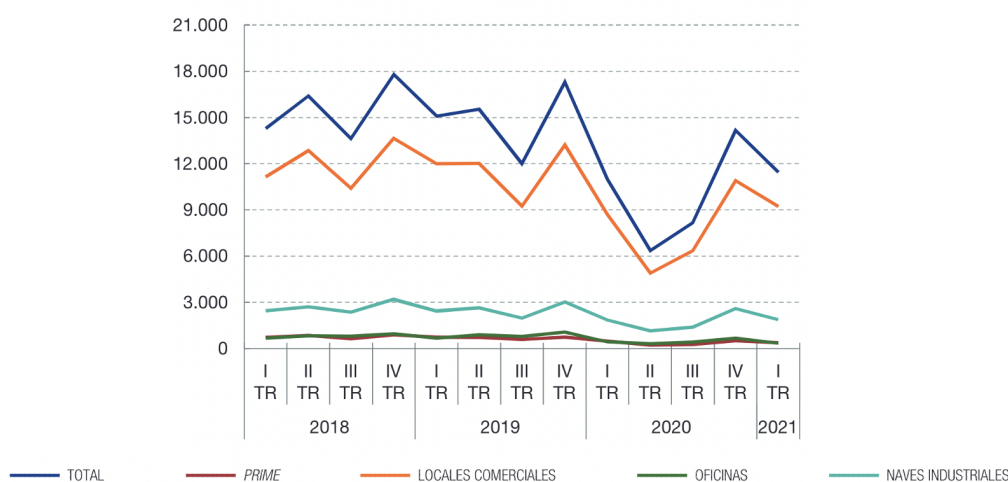
Esta disminución del número de transacciones relacionadas con inmuebles comerciales motivada por la pandemia no hace sino confirmar que el mercado inmobiliario comercial español es mucho más sensible a la coyuntura económica que el mercado de la vivienda (Fernández Cerezo et al., 2021) por lo que, en general, se vio más afectado por la reducción de la actividad del año 2020. Así pues, la demanda de oficinas, locales y naves industriales es más volátil y dependiente de las perspectivas económicas que la demanda de vivienda, mucho más estable.

Con todo, pese al reducido tamaño del mercado inmobiliario comercial sobre el total del mercado inmobiliario español y la mayor sensibilidad de este segmento a la situación macroeconómica, los inmuebles comerciales se erigen como una inversión muy interesante, pudiéndose alcanzar rentabilidades más altas que las que se obtendrían al invertir en una vivienda. Así se desprende, por ejemplo, de un estudio realizado por el portal inmobiliario Idealista (2020) según el cual, en el año previo a la pandemia, la rentabilidad bruta de la inversión en inmuebles comerciales, como locales y oficinas, superó a la de la inversión en vivienda. Mientras que la de los primeros se situó en el 8,9 % y el 9,9 %, respectivamente, la última se estancó en el 7,4 %.

Profundizando en la configuración del mercado inmobiliario en España, un análisis pormenorizado del sector de inmuebles de uso terciario puede revelar datos significativos. En primer lugar, resulta útil conocer el peso de cada tipología de inmueble en el mismo. La [Figura 1.2](#) deja patente el predominio de los locales comerciales en comparación con las naves industriales y, sobre todo, los edificios de oficinas y el segmento *prime*, esto es, aquellos locales situados en las zonas con mayor actividad comercial de las principales ciudades del país.

Teniendo en cuenta este hecho, Lamas y Romaniega (2022) apuntan la posibilidad de que los indicadores sobre la evolución del conjunto del sector inmobiliario comercial en España estén influenciados por la tendencia que siguen los locales comerciales. Por tanto, una corrección a la baja en el número de operaciones que se produjeron en el mercado inmobiliario comercial, como la que se muestra en la **Figura 1.1** durante 2020, puede contextualizarse con más detalle y comprenderse mejor tras una segmentación del sector según la clase de inmueble objeto de transacción.

En consecuencia, además de poder comprobar mediante la **Figura 1.2** la prevalencia de las operaciones de compraventa que involucran locales, también puede observarse por separado la evolución que sigue el mercado de cada tipo de inmueble de uso terciario. Principalmente, la tendencia es la misma en todos los casos. Si se analiza lo ocurrido durante 2020, la corrección debida a la pandemia experimentada por el conjunto del sector inmobiliario comercial español se tradujo en una reducción de las transacciones del 33,8 % (Fernández Cerezo et al., 2021).



**Figura 1.2:** Transacciones trimestrales en el sector inmobiliario comercial según el tipo de inmueble  
**Fuente:** Fernández Cerezo et al. (2021)

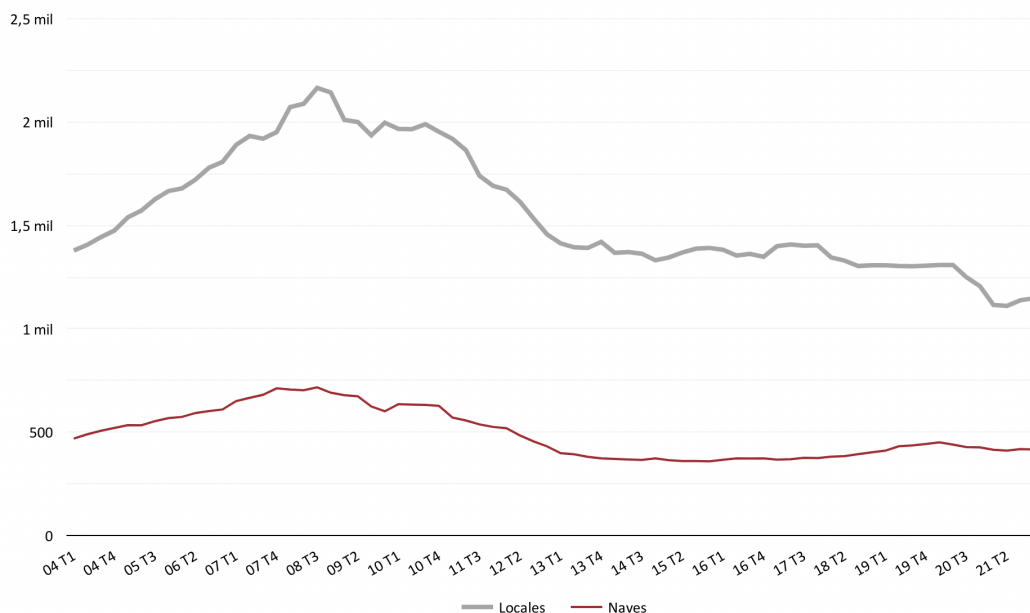
Los segmentos de oficinas y locales *prime* sufrieron retrocesos más significativos, del 47 % y el 47,9 %, respectivamente. En cambio, el descenso en el número de operaciones con naves industriales y locales comerciales fue similar en términos relativos. La caída interanual de las compraventas de las primeras fue de un 30,6 %, mientras que la minoración de las transacciones de los segundos alcanzó un 33,6 % (Fernández Cerezo et al., 2021). Estos datos reflejan que, por un lado, la reducción de las compraventas en el conjunto del mercado inmobiliario comercial español fue igual en magnitud a la de las compraventas de locales, confirmándose nuevamente el peso de este segmento en el sector. Sin embargo, por otra parte, durante la etapa de contracción económica derivada del estallido de la COVID-19, el segmento de naves industriales fue el que menos se resintió.

Si bien las malas perspectivas macroeconómicas, la intensidad de las medidas restrictivas de actividad y movilidad y la incertidumbre inicial en torno a la evolución de la pandemia desde el punto de vista sanitario provocaron, como en muchos otros sectores, un ajuste importante del

mercado inmobiliario comercial, la desigual caída relativa por segmentos puede explicarse a partir de la diferente afectación de las disposiciones de carácter restrictivo a cada uno de esos segmentos. Fundamentalmente, las medidas dirigidas al distanciamiento social perjudicaron especialmente a sectores como la hostelería, que hacen un uso intensivo de locales comerciales. En la misma línea, las limitaciones a la movilidad provocaron un fuerte descenso del comercio minorista, al mismo tiempo que se consolidó el crecimiento del comercio electrónico iniciado antes de la pandemia. Por último, la implementación obligatoria del teletrabajo en las empresas también contribuyó a reducir la demanda de oficinas.

Las medidas anteriores, en cambio, no tuvieron un impacto significativo en las actividades que suelen desarrollarse en naves industriales, sino que al contrario, impulsaron todavía más la necesidad de grandes centros logísticos distribuidos a lo largo de la geografía española para dar soporte al comercio *online* y mejorar la satisfacción del cliente en este canal. Además, en la [Figura 1.2](#) puede apreciarse otro punto a favor de las naves industriales, ya que en la fase de relajación de restricciones tras la primera oleada de la COVID-19 (segundo semestre de 2020, antes del empeoramiento de la situación sanitaria), las compraventas de estos inmuebles recuperaron niveles previos a la pandemia. Este fenómeno no se produjo en el resto de segmentos y apunta a una mayor estabilidad del mercado de naves en España a pesar de la coyuntura económica.

Un indicador alternativo al número de transacciones, a través del cual también puede evaluarse la evolución del mercado inmobiliario comercial español, es su índice de precios. La [Figura 1.3](#) muestra la variación del precio medio por metro cuadrado de locales comerciales y naves industriales. La tendencia de los precios en los últimos años refrenda las observaciones anteriores relativas a la mayor estabilidad del segmento de naves, dado que su precio medio no registra cambios sustanciales en comparación con los locales.



**Figura 1.3:** Evolución anual del precio medio de locales y naves en España (en €/m<sup>2</sup>)

**Fuente:** Fabra Garcés, Salvador Figueras y Gargallo Valero (2021)

Partiendo del hecho que ninguno de los dos segmentos ha recuperado el nivel de precios previo a la crisis económica de 2008, en el caso de las naves industriales sí se observa un leve crecimiento entre 2014 y 2019. En cambio, el precio medio de los locales comerciales presenta una ininterrumpida tendencia a la baja desde el último ciclo alcista. De hecho, durante 2021 los locales comerciales han registrado un mínimo histórico del precio medio, un ajuste interanual del -4,8 % que vino precedido por una caída del 7,6 % en 2020 (Fabra Garcés, Salvador Figueras & Gargallo Valero, 2021).

No obstante, el ajuste en el precio medio de las naves industriales ha sido más moderado, con un descenso en términos anuales del 2,3 % en 2021 y un 3,8 % en 2020. En resumen, el índice de precios evidencia, en el mismo sentido que los datos sobre compraventas, un mayor castigo al mercado español de locales durante la pandemia, pero según Fabra Garcés, Salvador Figueras y Gargallo Valero (2021) también sugiere que la dinámica de estos inmuebles anterior al *shock* económico de principios de 2020 ya era desfavorable en España en comparación con otros segmentos del sector.

## 1.2. Motivación

El mercado inmobiliario puede catalogarse como una interesante alternativa de inversión frente a los activos de renta fija. Los bienes inmuebles proporcionan una rentabilidad muy superior a, por ejemplo, los Bonos del Estado a 10 años, a pesar de que ambos constituyen inversiones que se asocian con un perfil de inversión conservador y están alejadas de inversiones sujetas a una mayor incertidumbre como los activos de renta variable. De acuerdo con un estudio de Idealista (2022b), los últimos datos anuales disponibles, referentes a 2021, indican que la rentabilidad bruta de la compra de una vivienda en España para su alquiler, que alcanza el 6,9 %, es 14 veces superior a la rentabilidad de los Bonos del Estado a 10 años en el mismo período, que se sitúa en el 0,5 % según el Banco de España (2022).

Además, el citado informe de Idealista revela que en el mercado inmobiliario comercial español la rentabilidad bruta es incluso mayor a la del mercado residencial, con tasas del 9,4 % en el caso de los locales comerciales y el 11,2 % en el segmento de oficinas. Por ello, a pesar de constituir un mercado más reducido y con una mayor propensión a entrar en un ciclo bajista si la coyuntura económica es desfavorable, la atención de los inversores en la actualidad parece centrarse en los bienes inmuebles de uso terciario.

Debido a la relevancia del sector inmobiliario en general y del segmento comercial en particular como mercado de inversión, es necesario el mayor número posible de indicadores para que futuros inversores puedan orientarse sobre la situación del mismo. Sin embargo, los organismos oficiales, como el Banco de España, analizan sobre todo el mercado residencial por su importancia más allá de la inversión y apenas ofrecen índices de precios relativos a los inmuebles comerciales. Además, el único indicador oficial sobre rentabilidad del alquiler aportado por el Banco de España corresponde al conjunto del mercado inmobiliario español.

En general, los estudios realizados por el sector privado, como los citados anteriormente, tampoco ofrecen una visión completa del mercado inmobiliario comercial, dado que no consideran el segundo segmento por volumen de operaciones, el de naves industriales, cuando en la actualidad es el que impulsa el mercado inmobiliario en España (Aranda, 2021). La consolidación del comercio

*online*, que potencia la demanda de naves industriales al mismo tiempo que reduce la de locales, y la estabilidad de las naves como activo, en tanto que son inmuebles demandados por sectores esenciales como el farmacéutico y el alimentario, auguran que la inversión en este segmento seguirá consolidándose y la rentabilidad esperada será atractiva para los inversores (Oi Realtor, 2021).

Por ello, conviene disponer de un indicador sobre la rentabilidad del alquiler de naves industriales para facilitar la toma de decisiones de los inversores, pero también es útil desde la perspectiva macroeconómica. Según Casas-Rosal, Núñez-Tabales y Caridad (2018) las ratios sobre la evolución del mercado inmobiliario comercial contribuyen a mejorar la radiografía de la situación económica en una zona determinada. Cuando dichas ratios son positivas, reflejan claramente una economía revitalizada y activa.

Una de las regiones con mayor actividad económica es la Comunidad Valenciana, en la que, además del turismo, tienen un peso significativo las industrias química y de la automoción o los sectores TIC y agroalimentario (Instituto de Comercio Exterior, s.f.). El dinamismo de la economía valenciana refrenda la afirmación anterior, puesto que el crecimiento de las compraventas en el mercado inmobiliario repuntaba a finales de 2020, aún en plena pandemia, gracias a los inmuebles comerciales (Cervellera, 2021), en especial las naves industriales. De hecho, durante 2021 la superficie transmitida de naves industriales en la Comunidad Valenciana fue la segunda mayor de España (Fabra Garcés, Salvador Figueras & Gargallo Valero, 2021), solo por detrás de Cataluña y por delante de importantes polos industriales como la Comunidad de Madrid o el País Vasco, motivando así el estudio de la inversión en este sector.

### 1.3. Objetivos

El objetivo principal del presente Trabajo Final de Grado consiste en calcular la ratio de rentabilidad bruta del alquiler de las naves industriales situadas en la Comunidad Valenciana a partir de una muestra de inmuebles de esta tipología en venta y en alquiler. Concretamente, se pretende hallar la rentabilidad bruta atendiendo a diferentes características de las naves industriales, de modo que pueda determinarse si existen diferencias en términos de rentabilidad cuando se consideran variables como la ubicación, la superficie o el equipamiento de las instalaciones.

A fin de lograr el objetivo principal, deben abordarse los objetivos específicos que se identifican a continuación:

- Conceptualización de las inversiones inmobiliarias y revisión de los indicadores de rentabilidad más habituales.
- Familiarización con el *web scraping* como método para la obtención de datos.
- Manejo del lenguaje de programación R, que constituye la base para el desarrollo del proyecto.
- Implementación de técnicas de *web scraping* para la construcción de una base de datos con ofertas de venta y alquiler de naves industriales en la Comunidad Valenciana.

- Aplicación de técnicas estadísticas para el análisis y la interpretación de los datos recopilados atendiendo a las diferentes variables según las cuales pueden clasificarse los inmuebles objeto de estudio.

En este sentido, a partir de los objetivos específicos descritos pueden diferenciarse dos partes. La primera se centra en la obtención de una muestra del mercado de naves industriales a partir de la cual pueda calcularse la ratio de rentabilidad bruta por metro cuadrado, siendo necesario elegir una fuente de información adecuada y desarrollar un mecanismo de acceso a la misma. Esta primera fase es imprescindible para poder desarrollar la segunda parte, que versa sobre el objetivo principal de este Trabajo Final de Grado.

## **1.4. Orden documental**

Este apartado detalla la organización por capítulos del contenido tratado en la presente memoria, de modo que pueda servir al lector como guía para facilitar la búsqueda de información a lo largo de la misma y lograr una mejor comprensión del trabajo. El resto del documento se organiza en torno a los siguientes capítulos:

### **Capítulo 2. Marco conceptual**

En este capítulo se describen dos conceptos básicos en el desarrollo del trabajo: desde la perspectiva económico-financiera, se revisa el concepto de inversión, con atención particular a las inversiones inmobiliarias y a los indicadores más empleados para medir la rentabilidad y, desde el lado técnico, se introduce el *web scraping* como método que habilita la recopilación de datos accesibles en Internet.

### **Capítulo 3. Metodología**

En este capítulo se explica el procedimiento que se ha seguido para desarrollar el trabajo. Paso a paso, se expone el método empleado para obtener información a partir de la cual se realiza el estudio definido como objetivo, el tratamiento previo de la información, el análisis estadístico y la forma de estimar la ratio de rentabilidad bruta del alquiler de naves industriales.

### **Capítulo 4. Resultados y discusión**

En este capítulo se exponen los índices de rentabilidad bruta del alquiler de naves industriales calculados y se comenta la variación de estas ratios en función de la zona geográfica considerada. También se repasa, en estas mismas localizaciones, el precio por metro cuadrado de los inmuebles utilizado en los cálculos.

### **Capítulo 5. Conclusiones y propuestas de trabajo futuro**

En este capítulo final se presentan las conclusiones alcanzadas tras la realización del trabajo y vinculadas a los objetivos establecidos inicialmente, además de ofrecer una valoración global del proyecto y plantear líneas futuras para su ampliación o aplicación en un ámbito distinto al desarrollado.





En este capítulo se profundiza en los elementos alrededor de los cuales se desarrolla el presente trabajo. Por un lado, se introduce el concepto de inversión en el ámbito económico-financiero y, más específicamente, se definen las inversiones inmobiliarias. En relación a las inversiones, se comenta la importancia de la rentabilidad como ratio asociada a las mismas, discutiendo diferentes criterios para su cálculo. Por otra parte, desde el punto de vista técnico, se realiza una breve presentación del *web scraping*, una metodología a través de la cual pueden recopilarse datos disponibles en Internet con el objetivo de que constituyan la base para la realización de diferentes estudios y análisis.

## 2.1. Rentabilidad en inversiones inmobiliarias

### 2.1.1. Concepto de inversión

Una persona o una empresa que dispone de recursos económicos puede elegir gastarlos, recibiendo a cambio una recompensa o satisfacción. Sin embargo, también puede optar por destinar sus recursos económicos a otros productos que, si bien no le ofrecerán satisfacción alguna en el presente, le permitirán recuperarlo junto con una compensación, suponiendo una mayor recompensa futura. Esta segunda operación recibe el nombre de inversión, mientras que la persona física o jurídica que la lleva a cabo es un inversor.

Singh (2016) define una inversión como un proceso en el que el inversor compromete un capital a largo plazo con el objetivo de conseguir ingresos adicionales a los fondos comprometidos, de forma que aumente su valor. Este crecimiento se conoce como rendimiento. Las inversiones, asimismo, pueden clasificarse en dos tipos, financieras y productivas.

Desde el punto de vista financiero, una inversión consiste en destinar recursos monetarios para obtener ingresos regulares futuros derivados de los mismos en forma de intereses, rentas, dividendos o plusvalías (apreciación del capital). En este sentido, en una inversión financiera lo importante no es el destino del capital vinculado, sino el rendimiento financiero que se puede obtener, reduciendo de este modo la inversión a meras transferencias de activos de un agente económico a otro (Singh, 2016). En las inversiones financieras este rendimiento, es decir, la consecución de mayores o menores ingresos, está ligado al riesgo de la inversión. En consecuencia, un mayor retorno de la inversión lleva aparejado un mayor riesgo, y viceversa, por lo que en esta clase de inversiones el inversor debe establecer una escala considerando el rendimiento que desea obtener y el riesgo que está dispuesto a asumir.

No obstante, los economistas trabajan con otro concepto de inversión. En este caso, se entiende por inversión la acción de una empresa consistente en destinar bienes y servicios para la producción de otros bienes y servicios (Singh, 2016). Supone, por tanto, la formación de capital productivo

nuevo (instalaciones, maquinaria, capital humano, etc.). En una inversión productiva, la empresa puede medir el rendimiento como el incremento de su capacidad de producción o la mejora de su posición en el mercado.

Una vez introducido el concepto de inversión y comentados los dos grandes grupos según los cuales puede catalogarse una inversión, es posible caracterizar una inversión inmobiliaria. En la Norma Internacional de Contabilidad (NIC) 40 sobre propiedades de inversión, las inversiones inmobiliarias o propiedades de inversión se definen como

[...] propiedades (terreno o un edificio —o parte de un edificio— o ambos) mantenidas [...] para obtener rentas o apreciación del capital o con ambos fines y no para:

- (a) su uso en la producción o suministro de bienes o servicios, o bien para fines administrativos;
- (b) su venta en el curso ordinario de las operaciones.

Junta de Normas Internacionales de Contabilidad (2018)

Esta definición presenta las inversiones inmobiliarias como inversiones financieras. La misma NIC también hace referencia a las inversiones productivas en bienes inmuebles, aunque no les otorga la etiqueta de inversión:

Propiedades ocupadas por el dueño son las propiedades que se tienen (por parte del dueño o por parte del arrendatario como un activo por derecho de uso) para su uso en la producción o suministro de bienes o servicios, o bien para fines administrativos.

Junta de Normas Internacionales de Contabilidad (2018)

Así pues, al hablar de inversiones inmobiliarias se hace referencia a inversiones de tipo financiero en las que se utiliza un bien inmueble para generar ingresos regulares e independientes, tal y como explica Singh (2016). En el caso de las propiedades de inversión, pueden esperarse de estas ingresos en forma de rentas, plusvalías o ambas. La NIC 40 presenta algunos casos de inversiones inmobiliarias, por ejemplo un terreno que se mantiene en propiedad para obtener apreciación en el largo plazo o un edificio desocupado que se destina al alquiler a través de uno o más arrendamientos operativos.

A partir de la conceptualización de las inversiones se constata que tras ellas subyace un flujo de cobros y pagos, siendo este una de los elementos principales para valorar una inversión, esto es, estudiar cuál resulta más interesante para el inversor. De acuerdo con Brotons Martínez (2017), los componentes principales de una inversión son los siguientes:

- Desembolso inicial: constituye el pago inicial para la adquisición de un activo.
- Vida útil: es el horizonte temporal de la inversión.
- Cobros: son los ingresos que suponen un flujo de entrada, esto es, que se producen en un determinado período y no están diferidos.
- Pagos: son los gastos que constituyen un desembolso efectivo de recursos monetarios.
- Flujos netos de caja: se obtienen como la diferencia entre cobros y pagos en cada uno de los períodos en los que se divide la vida útil de la inversión.

### 2.1.2. Concepto de rentabilidad

La valoración de las inversiones no es un procedimiento directo. Por un lado, la naturaleza de las inversiones difiere mucho de unas a otras y, con ello, los componentes enumerados anteriormente. Por otra parte, los flujos de caja asociados a las inversiones normalmente se producen en instantes temporales diferentes y pueden no ser constantes, de modo que una comparación exhaustiva no es directa. Pese a ello, existe un indicador que se limita a medir la capacidad de la inversión de generar beneficio. Dicho indicador es la rentabilidad, que indica la variación experimentada por una inversión durante un período de tiempo determinado (Brotons Martínez, 2017). Así pues, se calcula como la relación entre el beneficio que genera una inversión y los recursos monetarios comprometidos en el instante inicial. La rentabilidad es una ratio muy valiosa para estudiar el comportamiento de una inversión y poder establecer comparaciones con proyectos alternativos.

Cabe resaltar que, al definir una inversión, se ha hecho referencia a los ingresos regulares que se derivan de comprometer recursos monetarios en un instante inicial. Sin embargo, la definición de rentabilidad de Brotons Martínez (2017) tiene en cuenta el beneficio que obtiene el inversor. Esta dualidad se explica a partir de los componentes de una inversión ya que, como se ha visto, además de cobros también pueden existir pagos. La consideración o no de dichos pagos da lugar a dos variantes del beneficio, el beneficio bruto y el beneficio neto y, por tanto, a dos tipos de rentabilidad, la rentabilidad bruta y la rentabilidad neta (Raisin, s.f.).

La rentabilidad bruta solamente considera los ingresos que genera una inversión en relación al desembolso inicial y no se deducen otros gastos, mientras que se entiende por rentabilidad neta aquella que también considera estos gastos en los que se incurre durante la vida útil de la inversión. La rentabilidad bruta es una ratio más sencilla de obtener, en tanto que los gastos asociados a una inversión pueden ser difíciles de estimar dependiendo de la naturaleza de la misma. A cambio, la rentabilidad bruta no es una métrica de la rentabilidad real que puede esperar un inversor; para ello, es imprescindible calcular la rentabilidad neta.

En relación con la rentabilidad bruta y la rentabilidad neta, también puede distinguirse entre rentabilidad económica y rentabilidad financiera (Contreras, 2006). Por una parte, la rentabilidad económica (ROI, por sus siglas en inglés), es la razón entre el beneficio obtenido y el activo de la inversión, mientras que la rentabilidad financiera (ROE, por sus siglas en inglés) relaciona el beneficio con los fondos propios. De esta forma, el cálculo de la rentabilidad económica obvia la procedencia de los recursos monetarios para adquirir el activo, mientras que la rentabilidad financiera mide la capacidad de generar beneficio a partir de los fondos propios y no considera la financiación externa.

Asimismo, la rentabilidad puede particularizarse para las inversiones inmobiliarias. Retomando los ejemplos de inversión inmobiliaria descritos en la NIC 40, el beneficio que genera un bien inmueble que se tiene en propiedad sería la plusvalía que se obtendría en una futura venta. Esta plusvalía se identifica con la variación que experimenta el valor de la inversión, de acuerdo con la definición de Brotons Martínez (2017). En el caso de un inmueble que se arrienda a terceros, el beneficio proviene de las rentas del alquiler recibidas periódicamente.

Nuevamente, con esta información sobre los ingresos obtenidos únicamente puede hallarse la rentabilidad bruta, ya que la tenencia de un inmueble también genera una serie de gastos. El cálculo de la rentabilidad neta requeriría incorporar el pago del Impuesto sobre Bienes Inmuebles,

por ejemplo, además de otros gastos fijos. Adicionalmente, podría considerarse igualmente la financiación para adquirir inicialmente el inmueble, si la hubiere, en cuyo caso se estaría obteniendo una ratio de rentabilidad financiera neta.

Uno de los inconvenientes de la rentabilidad como métrica para la valoración de inversiones es que supone que el valor del dinero no varía con el paso del tiempo. Sin embargo, esta suposición no se corresponde con la realidad. Por este motivo, existen otros indicadores que sí consideran el descuento de los flujos de caja futuros al momento actual mediante una tasa de descuento adecuada. Son los llamados criterios dinámicos, en contraposición a los criterios estáticos (Brotons Martínez, 2017). Los criterios dinámicos solucionan la problemática en torno al valor temporal del dinero y permiten una comparación de inversiones más rigurosa. Los criterios dinámicos más empleados son el valor actual neto (VAN) y la tasa interna de retorno (TIR).

Sin embargo, la rentabilidad es un indicador que suele presentarse calculado en base a un horizonte temporal de un año. En estos casos, el problema del valor temporal del dinero se minimiza, de forma que la facilidad para calcular la ratio de rentabilidad e interpretar su significado convierten a este indicador en la métrica de referencia para la mayoría de inversores.

### 2.1.3. Ratio de rentabilidad bruta del alquiler

En el contexto de este trabajo, se persigue hallar la rentabilidad del alquiler de naves industriales, por lo que se analiza esta tipología de inversión inmobiliaria. En este sentido, se entiende por rentabilidad del alquiler de una inversión inmobiliaria la determinada por las rentas ingresadas en concepto de arrendamiento en relación con el desembolso inicial que supone la adquisición del inmueble. Dada la dificultad a la que se ha apuntado previamente para estimar los gastos fijos de explotación a los que debe hacerse frente por la propiedad del inmueble, únicamente se consideran los flujos de entrada, que se corresponden con las rentas ingresadas.

Por tanto, de acuerdo con estas puntualizaciones y con las clasificaciones anteriores, se pretende calcular una ratio de rentabilidad económica bruta en términos anuales, que puede expresarse matemáticamente en tanto por cien como

$$r_{\text{anual}}(\%) = \frac{\text{Renta anual alquiler}}{\text{Precio compra inmueble}} \cdot 100 \quad (2.1)$$

El indicador de rentabilidad bruta del alquiler calculado según la ecuación (2.1) sirve para determinar la rentabilidad de una inversión en particular. Sin embargo, la misma ratio puede generalizarse para que aporte información sobre la rentabilidad esperada de una inversión en el mercado inmobiliario, en este caso de naves industriales. Considerando el conjunto del mercado de naves, el valor inicial de la inversión puede estimarse a partir del precio de los inmuebles en venta, mientras que los ingresos anuales obtenidos en concepto de renta del alquiler pueden hallarse a partir de los precios de la oferta de inmuebles en alquiler. Además, para una mayor generalización del índice de rentabilidad, conviene normalizar ambos términos de la ecuación (2.1) por unidad de superficie, resultando en

$$RTP(\%) = \frac{\text{Precio anual/m}^2 \text{ alquiler}}{\text{Precio/m}^2 \text{ venta}} \cdot 100 \quad (2.2)$$

Este indicador de rentabilidad bruta del alquiler, que en el mundo anglosajón se conoce como RTP (rent-to-price), suele expresarse en términos medios o medianos y es la ratio cuya distribución quiere estimarse para el mercado de naves industriales de la Comunidad Valenciana.

## 2.2. Web scraping

La creciente digitalización ha provocado un aumento exponencial de los datos disponibles en Internet. Esta información se caracteriza fundamentalmente por ser heterogénea y dinámica, ya que los usuarios de la Web la generan y modifican de forma continuada. Esta heterogeneidad permite establecer diferentes clasificaciones de la información presente en Internet.

De acuerdo con Diouf et al. (2019), la información puede aparecer estructurada, semiestructurada o no seguir ningún tipo de estructura. Del mismo modo, puede clasificarse según la fuente de donde proceda. También puede distinguirse entre información cuantitativa o cualitativa, atendiendo a su naturaleza, o entre información útil o inútil, en función del valor que aporta dicha información. Estas diferencias en los datos accesibles a través de Internet aparecen en el momento de su creación, en tanto que los usuarios generan datos en diferentes formatos, por ejemplo, sitios web, bases de datos, correos electrónicos, publicaciones en redes sociales o contenidos multimedia, por citar los más populares.

La abundante cantidad de información, la permanente disponibilidad de la misma en la práctica y su calidad, si esta es seleccionada cuidadosamente, permite su utilización en todo tipo de estudios, tanto de investigación como comerciales, ofreciendo además muchas ventajas. Debido a la naturaleza de la información almacenada en la Web, los datos disponibles pueden considerarse una muestra en tiempo real de multitud de procesos y relaciones sociomateriales a partir de la que pueden desarrollarse análisis mucho más precisos, rigurosos y frecuentes (Krotov, Johnson & Silva, 2020). Además, al recurrir a la información disponible en Internet se simplifica el proceso tradicional de recopilación de datos, menos eficiente y más costoso, al mismo tiempo que se reduce la probabilidad de no obtener datos de ningún tipo sobre un determinado fenómeno.

Sin embargo, si bien en líneas generales la información de la Web puede obtenerse con mayor facilidad y rapidez que otras alternativas, deben diseñarse métodos potentes y eficientes que permitan tratar con las diferentes formas en las que puede presentarse dicha información para completar con éxito la recopilación de datos. En este sentido, el *web scraping* es una de las metodologías que gozan de mayor difusión y reconocimiento a la hora de recopilar datos procedentes de Internet.

### 2.2.1. Concepto de *web scraping*

El *web scraping* o extracción web puede definirse como una técnica para la extracción de datos disponibles en Internet y su almacenamiento informático, en forma de fichero o base de datos, con el objetivo de recuperarlos o analizarlos posteriormente (Zhao, 2022). Este procedimiento puede realizarlo un usuario manualmente o un robot de forma automatizada. En este segundo caso, el bot recibe el nombre de *web scraper* o rastreador web y consiste en un *software* que simula el comportamiento humano a la hora de navegar por Internet para recopilar información de

diferentes sitios web y estructurarla para ofrecérsela de forma ordenada al usuario (Diouf et al., 2019).

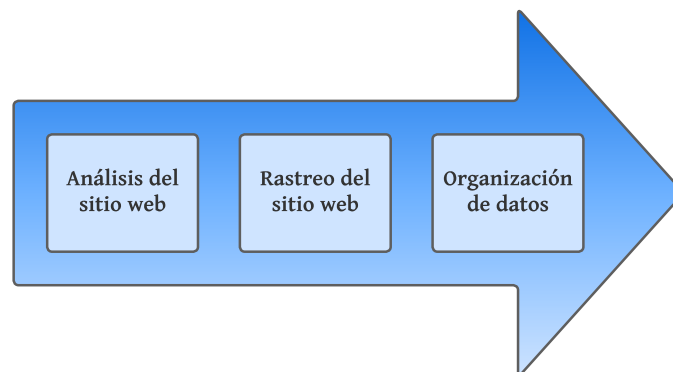
La definición anterior puede completarse a partir de la que proporcionan Glez-Peña et al. (2014), puesto que incide en una característica esencial de esta metodología que debe contemplarse tanto si se recurre a una extracción manual como automatizada: dicha extracción y el almacenamiento de los contenidos de interés obtenidos de Internet deben ser sistemáticos.

Pese a que según estas definiciones el *web scraping* puede ser manual o automatizado, mediante este término suele hacerse referencia a la extracción y organización automatizada de la información. Así pues, la técnica ha evolucionado desde procedimientos asistidos por humanos a sistemas completamente automatizados que se adaptan mejor a la gran variedad de escenarios y formatos disponibles en la Web (Zhao, 2022), además de tener una mayor capacidad para manejar datos y completar la tarea con mayor rapidez (Diouf et al., 2019).

Algunos de los usos más habituales del *web scraping* son el seguimiento y comparación de precios, la recopilación de reseñas de productos, la generación de bases de datos de anuncios inmobiliarios, el seguimiento de la información meteorológica o la detección de cambios de sitios web (Zhao, 2022). Por tanto, se justifica el empleo de un *web scraper* en un proyecto relacionado con el mercado inmobiliario como el que se desarrolla en este documento.

### 2.2.2. Procedimiento de extracción de datos

Krotov, Johnson y Silva (2020) establecen tres etapas interrelacionadas en el *web scraping*, que pueden identificarse en la Figura 2.1 y se derivan de las definiciones anteriores. Las tres etapas del *web scraping* son el análisis del sitio web, el rastreo del sitio web y la organización de datos.



**Figura 2.1:** Etapas del *web scraping*  
**Fuente:** Elaboración propia a partir de Krotov, Johnson y Silva (2020)

En primer lugar, la etapa de análisis del sitio web implica examinar un sitio web o un conjunto de ellos para comprender la estructura subyacente y el modo de almacenar la información. Esta etapa es la más analítica de todo el procedimiento y requiere que el usuario esté familiarizado con

las tecnologías web, principalmente la arquitectura web, los lenguajes de programación empleados en la construcción de las páginas web y las bases de datos.

Tras el análisis, la fase de rastreo del sitio web abarca el desarrollo y la ejecución de un bot que, de forma automática, navegue por un sitio web o un conjunto de ellos, analice la información disponible y recopile la que sea necesaria para el usuario o resulte de su interés. Por tanto, la acción del *web scraper* comienza en esta etapa y puede dividirse en dos módulos. Inicialmente, un programa de *web scraping* busca establecer conexión con el sitio web objetivo enviando una petición a través del protocolo HTTP (Glez-Peña et al., 2014), el mismo que se utiliza cuando un usuario cuando pretende visualizar en un navegador web cualquier página.

A continuación, cuando el servidor web ha procesado la petición HTTP del *web scraper*, la página web o recurso específico solicitado por el programa se envía de vuelta al mismo si no hay errores ni restricciones. Este paso se correspondería con la efectiva visualización de la página web por parte del usuario en su navegador. La segunda tarea del bot en la etapa de rastreo corresponde al análisis del documento enviado por el servidor, generalmente el código fuente del sitio web en formato HTML, para extraer la información buscada. Para ello, es común recurrir a la coincidencia, bien con expresiones literales, con los elementos que componen la estructura de una página web o con el lenguaje de estilo asociado a dichos elementos (Glez-Peña et al., 2014).

Por último, una vez analizado el sitio web y extraída la información del mismo, la etapa de organización de datos se centra en transformar los contenidos extraídos en una estructura que sea adecuada para su almacenamiento y con la que, tras su debido procesamiento, sea posible trabajar y resulte de utilidad para algún propósito.

En los últimos años han proliferado muchas herramientas de *web scraping* que facilitan el proceso a los usuarios que no tienen el conocimiento suficiente sobre cómo afrontar las etapas que se han descrito (Krotov, Johnson & Silva, 2020). Sin embargo, estas herramientas no siempre funcionan como espera el usuario y pueden fallar a la hora de extraer datos de un sitio web cuya estructura sea peculiar. Por ello, es más deseable familiarizarse con lenguajes de programación diseñados para la ciencia de datos, siendo R y Python los más populares, para implementar un *web scraper*. Estos lenguajes son los preferidos porque cuentan con librerías que facilitan el acceso a los sitios web, el análisis del código fuente de los mismos y la codificación de la información, suponiendo una curva de aprendizaje pequeña (Glez-Peña et al., 2014).

### 2.2.3. Ética y legalidad del *web scraping*

Finalmente, para completar la caracterización del *web scraping* no puede obviarse la dimensión ético-legal que rodea a esta práctica. En última instancia, se está obteniendo información de terceros almacenada en servidores cuyos administradores pueden oponerse a que se realicen copias de su sitio web o a soportar tráfico generado por *software* automatizado. Desde la perspectiva legal, no existe legislación específica sobre el *web scraping* como actividad en sí, sino que las normas relacionadas con el mismo se centran en el acceso ilegal al sitio web, la infracción de derechos de autor y el daño a los servidores (Krotov, Johnson & Silva, 2020).

Los sitios web cuentan normalmente con unos términos de uso o servicio en los que, si se prohíbe explícitamente el acceso automatizado, el responsable del *web scraper* podría estar



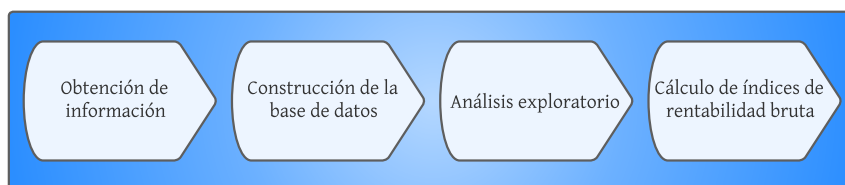
incurriendo en un acceso ilegal realizando peticiones al servidor web por esta vía. Sin embargo, según Zhao (2022) el propietario del bot podría refutar en un pleito la posición del administrador del sitio web argumentando que nunca se le presentaron dichos términos de servicio ni los aceptó. Igualmente, es difícil demostrar un quebrantamiento de los derechos de autor si el programa automatizado solo extrae información parcialmente y los datos de la web no están protegidos legalmente en su totalidad.

Un caso en el que puede demostrarse con más facilidad el perjuicio a un sitio web es aquel en el que las peticiones demasiado frecuentes de un *web scraper* a un servidor afectan a su funcionamiento normal e incluso lo dejan inoperativo, deshabilitando el acceso al contenido web. Técnicamente, este comportamiento recibe el nombre de ataque de denegación de servicio. No obstante, independientemente del recorrido legal en esta situación, los administradores de los sitios web suelen implementar medidas para detener o interferir la actividad de los bots y evitar el perjuicio derivado de tener que dejar de ofrecer el servicio a los usuarios humanos. Algunas de estas medidas son el análisis de la petición HTTP que recibe el servidor, puesto que las peticiones realizadas por usuarios humanos desde un navegador y las enviadas por *web scrapers* difieren, el bloqueo de direcciones IP sospechosas y los CAPTCHA, que son pruebas de desafío-respuesta para detectar programas automatizados.

A pesar de las lagunas legales y de la dificultad para justificar daños a raíz del *web scraping*, es deseable que los usuarios de esta técnica desarrollen herramientas acordes con la ética. Con este propósito, Krotov, Johnson y Silva (2020) enumeran algunas preguntas que los desarrolladores de *web scrapers* pueden plantearse para saber si están implementando un proyecto ético. Las más destacadas guardan relación con el respeto a los términos de servicio y los derechos de autor y el uso que se le pretende dar a la información extraída.

Así pues, debe tenerse en cuenta si dicha información es confidencial o puede comprometer de algún modo la privacidad del propietario del sitio web o terceros relacionados, pero también que la información obtenida sea de calidad y no se utilice para inducir la toma de decisiones erróneas por parte de los receptores de la misma. Al mismo tiempo, un *web scraper* ético debe detener su funcionamiento si el sitio web lo ha requerido por algún medio (las medidas descritas en el párrafo anterior) o si impacta significativamente en el valor del servicio ofrecido por el sitio web (por ejemplo, las visitas generadas por bots no pueden monetizarse).

Una vez introducidos los objetivos del proyecto y el contexto en el que se puede enmarcar, en este capítulo se presenta la metodología que se ha seguido en el presente trabajo. La metodología empleada consta de cuatro fases, representadas en la [Figura 3.1](#), desarrollándose todas ellas mediante el lenguaje de programación R, uno de los entornos *software* más utilizados en el campo de la estadística y la ciencia de datos. R es un lenguaje en constante evolución que cuenta con muchas librerías (extensiones) que simplifican el tratamiento de datos, los cálculos y la visualización de la información.



**Figura 3.1:** Metodología de trabajo

El primer paso del trabajo consiste en hallar información sobre la oferta de naves industriales en la Comunidad Valenciana y extraer un conjunto de datos que se utilizarán como muestra representativa del mercado, a partir de la cual se calculará la rentabilidad del alquiler en dicho sector. Para ello, se emplea la técnica de *web scraping* descrita en el [Capítulo 2](#), siendo necesario un análisis previo para determinar la fuente de información más adecuada. Esta fase se identifica como *Obtención de información* en la [Figura 3.1](#).

A continuación, la información bruta, es decir, los datos extraídos de Internet, deben ser preprocesados para garantizar el correcto desarrollo de un estudio, al mismo tiempo que se garantiza su reproducibilidad. Por tanto, en esta fase de *Construcción de la base de datos* se estructura la información obtenida para conseguir una representación ordenada de la misma. Este paso, de acuerdo con la [Figura 2.1](#), constituye realmente la última etapa del proceso de *web scraping*. No obstante, además de codificar y estructurar convenientemente la información, en este punto también se profundiza en los datos recopilados para tratar de obtener información adicional a partir de estos, recurriendo para ello a la técnica del *text mining* o minado de texto. Por último, para conseguir una base de datos representativa del mercado de naves, esta debe depurarse eliminando las observaciones anómalas y duplicadas que podrían influir en el estudio posterior.

Posteriormente, habiendo consolidado la información para la determinación de la rentabilidad bruta, el siguiente paso es analizar las distintas variables que describen los inmuebles de la muestra para caracterizar la oferta en el mercado de naves industriales de la Comunidad Valenciana. A

tal fin, se emplean técnicas de estadística descriptiva e inferencial en un proceso que se conoce como análisis exploratorio, tal y como recoge la [Figura 3.1](#). Finalmente, tras este estudio se está en disposición de hallar los índices de rentabilidad bruta del alquiler en función de las variables más interesantes que se desprendan del análisis exploratorio.

## 3.1. Obtención de información

### 3.1.1. Elección de la fuente de información

La confección de cualquier estudio requiere encontrar una fuente de información fiable y que cuente con suficientes datos recopilados para alimentar su elaboración. Tradicionalmente, la fuente de información por excelencia para la mayoría de analistas han sido los organismos públicos, en tanto que la información que ofrecen es accesible para toda la población y su recopilación se realiza a través de técnicas contrastadas. No obstante, como ya se ha comentado en el [Capítulo 1](#), las fuentes públicas en España presentan un gran inconveniente cuando se trata de estudiar el mercado inmobiliario, ya que dirigen su atención al sector residencial.

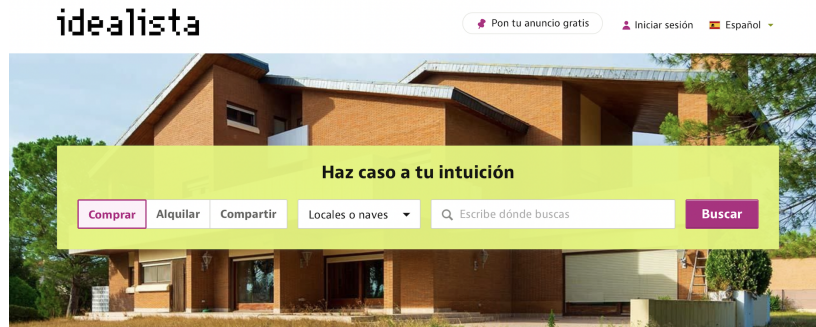
En cambio, contemplar fuentes privadas para el análisis del mercado inmobiliario comercial puede resultar de interés. En este sentido, los agentes inmobiliarios disponen de información sobre las transacciones inmobiliarias por su condición de intermediarios en estas operaciones, pero su ámbito de actuación generalmente se circunscribe a zonas geográficas limitadas. Sin embargo, en los últimos años se han popularizado los portales inmobiliarios en Internet y tanto particulares como agencias inmobiliarias tienen presencia en los mismos para ofertar inmuebles. Debido a la facilidad en el acceso y la navegación por estos portales, así como la inmediatez que ofrecen a la hora de consultar los anuncios, los portales inmobiliarios reciben mucho tráfico de usuarios y, como consecuencia, disponen de una oferta considerable de inmuebles tanto residenciales como comerciales repartidos por toda la geografía.

Entre las plataformas de búsqueda de inmuebles disponibles en España, Idealista es el portal web más visitado, con 54,1 millones de accesos en el mes de julio de 2022 (SimilarWeb, 2022). Este dato coloca a Idealista a mucha distancia de su principal competidor, Fotocasa, que recibe una cuarta parte de las visitas mensuales. Asimismo, Idealista cuenta con el mayor número de anuncios publicados, con todo tipo de inmuebles: viviendas, garajes, trasteros, locales comerciales, oficinas, naves industriales y terrenos. Por su posición como web líder en cuanto a transacciones inmobiliarias, muchos estudios sobre el mercado inmobiliario español utilizan este portal como fuente para obtener información, por ejemplo el trabajo de Casas-Rosal, Núñez-Tabales y Caridad (2018).

Por tanto, se decide justificadamente que la base para el estudio de la rentabilidad del alquiler en el mercado de naves industriales de la Comunidad Valenciana será la oferta de este tipo de inmuebles disponible en este portal web. Una vez escogida la fuente de información, el primer paso para llevar a cabo el proceso de *web scraping* consiste en analizar cómo se organizan internamente los anuncios.

Al acceder al sitio web de Idealista, el usuario visualiza la página de inicio, tal y como se observa en la [Figura 3.2](#). Para realizar una búsqueda, en esta página se elige si recuperar anuncios

de venta o de alquiler de inmuebles, así como la tipología de estos. Un primer vistazo a la página principal revela que la plataforma agrupa locales comerciales y naves industriales, por lo que más adelante deberá prestarse atención a los filtros disponibles que permitan diferenciar ambas categorías de inmuebles comerciales. El último filtro requerido para iniciar una búsqueda de anuncios es la ubicación, admitiendo desde barrios hasta provincias enteras, de forma que se dispone de una gran granularidad geográfica.

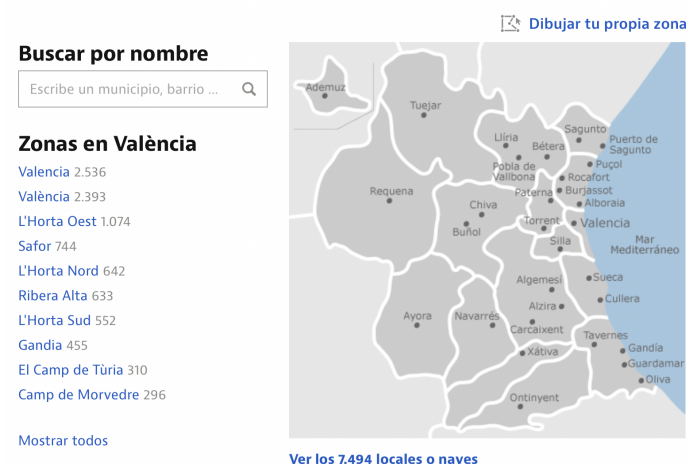


**Figura 3.2:** Página principal de Idealista  
**Fuente:** Idealista (2022a)

Tras definir los criterios de la búsqueda inicial, el sitio web presenta los resultados. Si se ha realizado una búsqueda sobre un barrio determinado de un municipio (máximo nivel de granularidad), el servidor devuelve directamente la lista de inmuebles ofertados en dicha área. Sin embargo, si la búsqueda se acota por zonas geográficas más grandes, aparece una página web con un mapa de la zona seleccionada como el de la [Figura 3.3](#). Este mapa siempre muestra el siguiente nivel de granularidad al indicado como filtro en la búsqueda inicial.

### Mapa de València provincia: locales o naves en venta,

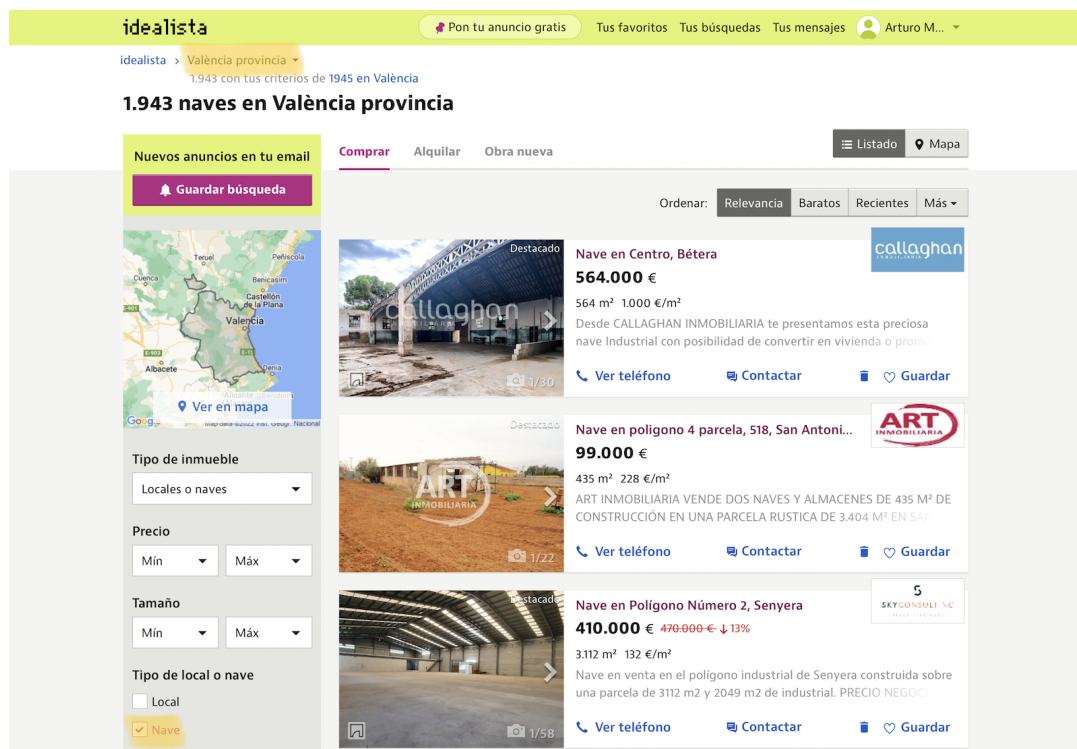
[Ver los 7.494 locales o naves](#)



**Figura 3.3:** Mapa de la provincia de Valencia en Idealista  
**Fuente:** Idealista (2022a)

Así, la **Figura 3.3** corresponde a una búsqueda de locales o naves en venta en la provincia de Valencia (mínima granularidad), por lo que el sitio web devuelve un mapa con el siguiente nivel de división territorial en la Comunidad Valenciana, el comarcal. En este punto, se sigue comprobando que la plataforma todavía no permite diferenciar entre locales comerciales y naves industriales. A continuación, tras ofrecer al usuario la posibilidad de modificar el área de búsqueda de inmuebles, el sitio web devuelve el listado con los anuncios correspondientes a la zona geográfica seleccionada. En esta fase de la búsqueda, Idealista presenta nuevos filtros relativos al precio, la superficie del inmueble o la fecha de publicación del anuncio, por ejemplo. No obstante, el más interesante en este caso es el que permite diferenciar entre locales y naves.

De este modo, es posible eliminar de entre los resultados devueltos por el sitio web gran parte de la oferta de locales comerciales, aunque el filtro no es infalible. En última instancia, son los usuarios los que suben los anuncios, y un análisis del listado revela que una parte importante de los mismos no especifica el tipo de inmueble del que se trata. Por ello, pese al filtro, estos anuncios aparecen con la etiqueta *Local o nave* en el título. Habiendo aplicado el único filtro de interés en el contexto de este proyecto (no se desea limitar a la hora de recopilar información ninguna variable), el listado de anuncios de la zona geográfica escogida tiene el aspecto que se observa en la **Figura 3.4**.



**Figura 3.4:** Listado de anuncios de naves en venta en Idealista en la provincia de Valencia

**Fuente:** Idealista (2022a)

Acerca de esta lista, cabe destacar que la plataforma presenta los anuncios divididos en páginas, a razón de 30 inmuebles por página. Sin embargo, el sitio web genera como máximo 60 páginas por búsqueda, es decir, el usuario puede consultar un total de 1800 anuncios en cada búsqueda.

Esto supone que, en aquellos casos en los que el total de anuncios registrados en Idealista en una determinada zona geográfica supera dicho número, no se muestran todos los resultados disponibles. Por ello, es importante acotar adecuadamente el filtro de ubicación con el objetivo de evitar la pérdida de inmuebles, como ocurriría en el caso que ilustra la [Figura 3.4](#), donde se registran 1941 naves en venta en la provincia de Valencia.

El último nivel del sitio web de Idealista es el formado por las páginas web de cada inmueble anunciado, en las que se describen con más detalle sus características y se proporciona la información de contacto de los propietarios o agentes inmobiliarios. En la página del anuncio de un inmueble, además del precio en euros, la superficie en metros cuadrados y la ubicación según los diferentes niveles de división territorial (barrio, distrito, municipio, comarca y provincia), también puede consultarse la descripción de la propiedad aportada por el usuario, así como fotografías y una lista parcialmente estandarizada de características relativas al equipamiento, la antigüedad y la situación general del inmueble.

### 3.1.2. Implementación del *web scraper*

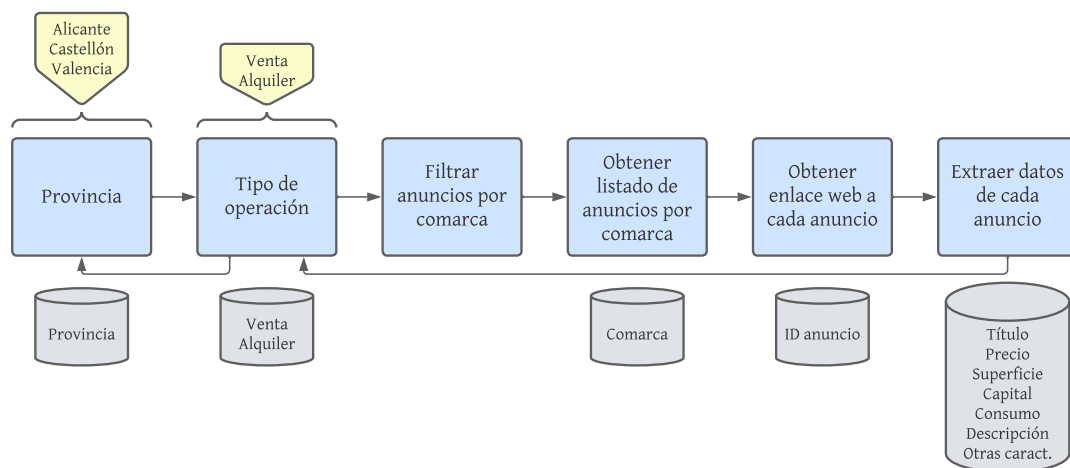
Habiendo comprendido la estructura del sitio web de Idealista, es posible diseñar un bot para extraer de forma automática la información de los anuncios de interés disponibles. Se opta por implementar un *web scraper* desde cero por la versatilidad que ofrece esta opción, al disponer de control absoluto sobre la forma en la que el *software* navega por las distintas páginas web. Además, el desarrollo de esta herramienta ha permitido profundizar en el manejo del lenguaje de programación R, sobre el que se implementa en su totalidad.

Desde el punto de vista técnico, el *web scraper* se basa en las librerías *httr* y *rvest* disponibles en R. La primera implementa métodos para interactuar con un sitio web a través del protocolo HTTP. De todos ellos, se utiliza fundamentalmente el método GET para recuperar una página web del servidor. Asimismo, *httr* posibilita configurar diferentes parámetros de conexión y personalizar las peticiones enviadas por el bot a través del método anterior para asemejar su presentación y comportamiento al de un usuario humano.

En cuanto a *rvest*, esta librería incluye las funciones necesarias para poder trabajar en R con el código fuente en formato HTML de las páginas web que devuelve el servidor y extraer datos de las mismas. Para ello, se filtra dicho código fuente por medio de la coincidencia literal con cadenas de caracteres o a través de la etiqueta HTML de un determinado elemento de la estructura de la página web en el que se ha identificado algún dato de interés. En consecuencia, otras librerías útiles son *stringr* para manipular cadenas de caracteres y *readr* para el mismo propósito, pero con los valores numéricos.

Por otra parte, el *web scraper* que se ha implementado aprovecha el soporte de R a la paralelización de procesos. A grandes rasgos, la paralelización consiste en repartir la carga de trabajo generada por un *software* determinado entre todos los núcleos del procesador de un ordenador, permitiendo completar partes de dicho trabajo al mismo tiempo, lo que redundaría en un menor tiempo de ejecución. En este contexto, la paralelización habilita una extracción más rápida de los datos del sitio web de Idealista.

Una vez explicadas las características más reseñables del bot implementado, conviene detallar también su funcionalidad. En este sentido, cabe destacar los datos que se extraen con el mismo y, en relación con ello, el comportamiento que sigue a la hora de realizar peticiones al sitio web de Idealista. La [Figura 3.5](#) pretende ilustrar el funcionamiento del *web scraper*, donde los bloques azules indican el orden que sigue el *software* mientras navega por el sitio web de Idealista y los bloques grises señalan los datos que se obtienen de cada anuncio y el instante de la ejecución del bot en el que se extraen y almacenan.



**Figura 3.5:** Diagrama de bloques del *web scraper* implementado para recopilar información de Idealista

Durante el estudio de la estructura de la plataforma se ha comprobado que existen diferentes niveles a la hora de filtrar los inmuebles por su localización, desde barrios de municipios (el nivel más particular) a provincias enteras (el nivel más general). Por ello, teniendo en cuenta que el análisis presentado en este trabajo abarca el territorio de la Comunidad Valenciana, se considera adecuada la división comarcal, en tanto que una división por municipios o más limitada supondría tener una muestra de pocas o incluso ninguna nave industrial en alguna de las zonas.

Justificado el nivel de división geográfica, el funcionamiento del bot desde que se inicia su ejecución hasta que finaliza la extracción de datos puede resumirse en los siguientes pasos:

1. Se obtiene la dirección web (URL) de Idealista correspondiente a los locales y naves en venta o alquiler de una provincia determinada.
2. Se añade al final de la URL anterior la etiqueta que indica que el usuario ha activado el filtro de naves industriales en la búsqueda, como se observa en la [Figura 3.4](#).
3. A partir del listado de anuncios de naves toda la provincia, se obtienen las URL de los listados de anuncios por comarca contenidas en el desplegable señalado en la [Figura 3.4](#).
4. En función del número de anuncios registrados en cada comarca, se generan automáticamente las direcciones web de cada una de las páginas de los diferentes listados añadiendo al final de la URL de la página inicial de cada listado el número de página.

5. Se agrupan las URL de todas las páginas con anuncios de naves que deben explorarse y se obtienen los enlaces de todos los anuncios.
6. Se descarga la página web de cada anuncio y se extrae la información pertinente detallada en la [Figura 3.5](#).
7. Se repite el proceso hasta que se han extraído los datos de todas las naves en venta y en alquiler de las tres provincias de la Comunidad Valenciana.

Siguiendo esta metodología, es posible garantizar que con este nivel de granularidad en ningún caso se superan los 1800 anuncios en la zona geográfica considerada (comarca), al mismo tiempo que se puede almacenar información sobre los inmuebles durante todo el procedimiento. Así pues, se conoce la provincia en la que se ubica la nave y si está destinada a la venta o el alquiler desde el principio. Es decir, en una iteración de la ejecución del *web scraper* todas las naves comparten estos atributos. A continuación, al filtrar por comarca se consigna este parámetro a todos los anuncios que aparecen en los listados correspondientes. Tras ello, al obtener las direcciones web de todos los anuncios se utiliza el mismo código numérico que tienen asignado en el sitio web de Idealista como variable para su identificación en la base de datos. Por último, el grueso de atributos de las naves se recopila de la página web del anuncio.

A modo de resumen, la [Tabla 3.1](#) recoge los datos sobre cada anuncio que se han extraído del sitio web de Idealista y su significado.

**Tabla 3.1:** Datos recopilados en el proceso de *web scraping*

Atributo	Descripción
ID	Código numérico para identificar de forma unívoca una nave
Tipo de anuncio	Indica si la nave se anuncia para su venta o alquiler
Título	Título del anuncio según aparece en la página
Precio	Precio de compra de la nave (venta) o renta mensual (alquiler) en euros
Superficie	Superficie en metros cuadrados de la nave
Provincia	Provincia en la que se ubica la nave
Comarca	Comarca en la que se ubica la nave
Capital	Indica si la nave se ubica en una capital de provincia (o en la ciudad de Elche)
Características	Lista con algunos atributos parcialmente estandarizados en todos los anuncios
Consumo	Etiqueta de eficiencia energética (A-G) de la nave
Descripción	Campo de descripción del anuncio de la nave

### 3.2. Construcción de la base de datos

Mediante la herramienta de *web scraping* implementada en este trabajo, el día 16 de agosto de 2022 se obtuvo información de un total de 5014 anuncios de inmuebles etiquetados como naves industriales en Idealista. A partir de estos datos, se construye una base de datos correctamente estructurada de forma tabular, en la que las filas son observaciones y las columnas medidas. En este caso, las observaciones se corresponden con la muestra de naves industriales extraída del sitio web, mientras que las columnas se corresponden con los diferentes atributos asociados a estos inmuebles, descritos en la [Tabla 3.1](#).



### 3.2.1. Generación de nuevas variables

En la lista de atributos se observan dos clases de datos que no ofrecen ninguna medida y que, por tanto, no pueden formar parte de la base de datos final sin un tratamiento previo. Se trata de los campos de descripción y de características básicas, que no se pueden identificar de forma directa como variables descriptivas de las naves industriales. Ambos atributos consisten en cadenas de texto (no estructuradas y semiestructuradas, respectivamente) que se pueden explorar aplicando metodologías adecuadas para encontrar patrones relevantes que podrían resultar de interés.

El *text mining* o minería de textos puede definirse como la técnica derivada del *data mining* o minería de datos consistente en aplicar una serie de métodos y algoritmos con el objetivo de encontrar patrones útiles y extraer información de textos no estructurados (Hotho, Nürnberger & Paaß, 2005). En el contexto de este trabajo, la minería de textos es una metodología que puede aplicarse al campo de descripción de los anuncios de naves industriales para tratar de encontrar variables descriptivas adicionales.

El preprocesamiento de textos consta de varias etapas. La primera de ellas es la *tokenización*, un proceso a través del cual se separan todas las palabras de un documento de texto eliminando todos los signos de puntuación, espacios y otros caracteres no imprimibles (Hassler & Fliedl, 2006). Adicionalmente, para lograr una mayor homogeneidad del texto en aras de un análisis posterior más sencillo, durante el proceso de *tokenización* también se completan las siguientes tareas:

- Se cambian todas las letras a minúsculas para eliminar la variabilidad en cuanto al lugar en el texto en el que aparecen las palabras, o por si algún usuario ha redactado la descripción del anuncio utilizando únicamente letras mayúsculas.
- Se excluyen todos los valores numéricos ya que el preprocesamiento del campo de descripción llevado a cabo se centra exclusivamente en el texto, dada la complejidad que conlleva el tratamiento de números.
- Se quitan las tildes de las vocales con el mismo propósito con el que se reducen las letras a minúsculas, una mayor uniformidad del texto.

Tras la *tokenización* se construye el diccionario, que registra las palabras únicas que aparecen en el conjunto de descripciones. Sobre este diccionario pueden aplicarse diversos métodos para reducir su tamaño. Uno de estos métodos es el filtrado, con el que se eliminan palabras del diccionario y, en consecuencia, de las descripciones. El filtro más habitual se basa en la detección de las palabras vacías (Yao & Ze-wen, 2011). Las palabras vacías son aquellos artículos, pronombres, preposiciones, conjunciones, etc. que carecen de significado por sí mismos, por lo que no aportan valor alguno al significado del texto. Así pues, el diccionario generado previamente se modifica eliminando las palabras vacías del diccionario español.

En este punto se cuenta con un diccionario que se puede examinar para determinar cuáles son las palabras más repetidas en las descripciones de los anuncios de naves industriales que se han extraído de Idealista. En la [Figura 3.6](#) están representados los 100 términos más frecuentes en los textos de descripción y en la que, a mayor tamaño de la palabra, mayor frecuencia de la misma. Por un lado, la nube de palabras ilustra el principal problema detectado durante la obtención de información mediante *web scraping*: la notable presencia de anuncios de locales comerciales en la

muestra, puesto que la palabra *local* aparece con la misma frecuencia que *nave*. En consecuencia, este inconveniente deberá tratarse posteriormente.



**Figura 3.6:** Nube de palabras con los 100 términos más repetidos en las descripciones de los anuncios

Por otra parte, la nube de palabras revela otros términos interesantes además de los esperados, siendo este el objetivo principal de aplicar el proceso de *text mining* a las descripciones. Dichas palabras aparecen en la [Figura 3.6](#) con diferentes tonalidades de azul. Los términos identificados que se considera que completan la caracterización de una nave industrial son los siguientes:

- Carretera. La presencia de esta palabra en la descripción puede significar que la nave industrial está bien comunicada, al ubicarse en las proximidades de carreteras principales.
- Diáfana. Una nave diáfana es aquella que carece de obstáculos y no está dividida en estancias, permitiendo destinar su uso a diversas actividades.
- Polígono. Con el uso de esta palabra en la descripción, el anunciante puede remarcar que la nave se ubica en un polígono industrial.
- Aparcamiento. Si este término aparece en la descripción, puede indicar que la nave cuenta con una zona para el estacionamiento de vehículos.

Aisladas las palabras más útiles, se aplica un último método para simplificar el diccionario generado previamente. Se trata del algoritmo de *stemming* de Porter (1980), diseñado para extraer el morfema o raíz de las palabras y evitar que la adición de sufijos para formar términos derivados penalice la frecuencia de aparición de estos morfemas, que son los que en última instancia dotan de significado a las palabras. Tras esta modificación del diccionario, se buscan en las descripciones los morfemas de las palabras enumeradas anteriormente y se considera que, si aparecen en el texto del anuncio, las naves correspondientes presentan dichos atributos.

Una vez finalizado el procesamiento de las descripciones, se lleva a cabo un procedimiento similar con la lista de características básicas de cada anuncio. No obstante, estos textos están

semiestructurados y su tratamiento es más sencillo. Se identifican como variables de interés el estado de la nave industrial (a estrenar, bueno y para reformar), el número de estancias (variable que se combina con la relativa a si la nave es diáfana definida previamente), el número de aseos, el año de construcción y el número de plantas.

Para determinar el valor de estas variables, se busca por coincidencia la presencia de las cadenas de caracteres en las que aparecen dichas variables y, en caso afirmativo, se procesan. Por ejemplo, si en la lista de características de un anuncio una de las entradas incluye la palabra *aseo*, dicha cadena de texto contendrá el número de aseos, que se extrae y se asigna como medida de esta variable a la nave correspondiente.

### 3.2.2. Limpieza de datos

A la hora de construir una base de datos válida para su análisis estadístico, otra de las etapas imprescindibles que debe completarse, además de la adecuada estructuración de la información, es el tratamiento de los datos recopilados con el objetivo de detectar la presencia de observaciones duplicadas u observaciones cuyas variables no tienen un valor asignado o registran valores anómalos.

#### Valores duplicados y faltantes

La supresión de observaciones duplicadas o de las que presentan valores faltantes de las medidas tiene como objetivo, por un lado, evitar desarrollar estudios posteriores sobre información redundante que podría afectar a los resultados que se obtengan en términos estadísticos. Por el otro, se consigue filtrar aquellas observaciones cuya utilización en esos estudios no será posible por la falta de concreción en las medidas.

En primer lugar, tienen la consideración de observaciones con valores faltantes aquellos anuncios que no contemplan si se trata de una nave ofertada para la venta o el alquiler, el precio de venta o el importe mensual del alquiler, la superficie del inmueble y su ubicación (provincia y comarca). Se considera que estas variables son las mínimas imprescindibles para poder caracterizar una nave. En total, existen 4 inmuebles que se eliminan de la muestra puesto que no consta su superficie. Se trata de anuncios promocionados en Idealista en los que los datos se presentan con otra estructura, de modo que el *web scraper* no ha podido recuperar esta información.

Seguidamente, se consideran observaciones duplicadas las correspondientes a anuncios en los que se repiten los mismos valores en cada una de las variables consideradas anteriormente, a las que se suma el título del anuncio. Se utiliza la información del título del anuncio para evitar descartar indebidamente anuncios de naves de la misma superficie y precio situadas en la misma comarca, pero en localidades diferentes. Así, tras aplicar este criterio para hallar los valores duplicados, se suprimen 168 entradas de la base de datos, un 3,4 % de la base de datos original. Generalmente, los valores duplicados se deben a que el mismo inmueble es publicado por varios anunciantes, por ejemplo, agencias inmobiliarias distintas o un particular y una agencia inmobiliaria.

### Valores anómalos

De igual manera que con los datos duplicados o faltantes, es importante detectar los valores anómalos de las observaciones. Rousseeuw y Hubert (2011) definen los *outliers* o valores anómalos como aquellas observaciones que difieren de la mayoría de la muestra. Pueden deberse a errores, a un registro incorrecto de las medidas o a que sean observaciones de características excepcionales. En consecuencia, estos registros no se ajustan adecuadamente a la muestra y conviene eliminarlos, ya que podrían impactar significativamente en los resultados y conclusiones posteriores si se mantienen de cara al análisis.

En el apartado anterior se ha constatado la notable presencia de locales comerciales en la muestra, por lo que queda pendiente tratar con los mismos. De acuerdo con la definición dada, estos inmuebles se pueden considerar observaciones anómalas y, por tanto, deben descartarse para que la muestra final sea representativa del mercado de naves industriales. Como es difícil definir un criterio preciso para la detección de los locales, se opta por eliminar de la base de datos todos aquellos anuncios cuyo título contiene la palabra *local*, o en cuya descripción también aparece dicho término y no está presente la palabra *nave*. Aún a costa de una excesiva rigurosidad, de esta manera se excluyen 2427 inmuebles de la muestra, un 48,4 % de la muestra original.

Una vez que se le ha dado solución al problema de la presencia de locales comerciales en la muestra, a continuación pueden aplicarse diversas técnicas estadísticas para la detección de *outliers*, las cuales se agrupan según el enfoque. En este sentido, se distinguen los enfoques univariante y multivariante (Leys et al., 2019). Los métodos de detección univariante suponen que una observación es anómala cuando registra un valor anormal en una única dimensión (variable). Para ello, se utiliza un indicador robusto de tendencia central de la muestra y se concluye que una observación es anómala cuando se aleja mucho de dicho indicador.

En cambio, las técnicas de detección multivariante consideran que una observación difiere significativamente de la mayoría de la muestra cuando se han tenido en cuenta dos o más variables. En este caso, al considerar  $n$  dimensiones, se etiquetan como observaciones anómalas aquellas que están muy alejadas de una figura  $n$ -dimensional. Por ejemplo, cuando se aplican métodos multivariante con dos variables, la figura bidimensional de referencia que se genera y engloba los valores normales, dejando los valores extremos fuera de su perímetro, es una elipse.

Dentro del enfoque univariante, uno de los estadísticos robustos más explorados es el rango intercuartílico o IQR (Rousseeuw & Hubert, 2011). Considerando una variable determinada, el rango intercuartílico se define como la diferencia entre el tercer y el primer cuartil de los valores de dicha variable en la muestra. En base al rango intercuartílico, matemáticamente puede determinarse que un valor es anómalo cuando se encuentra fuera del intervalo dado por

$$I = [q_1 - b \cdot IQR; q_3 + b \cdot IQR], \quad (3.1)$$

donde  $q_1$  y  $q_3$  denotan el primer y el tercer cuartil, respectivamente, y  $b$  es una constante que puede tomar los valores 1,5 y 3 en función de la rigurosidad que se desee a la hora de eliminar valores anómalos.

En cuanto al enfoque multivariante, un estadístico ampliamente empleado es la distancia de Mahalanobis, una medida de la distancia entre un punto de una distribución multivariante y

su media (Maesschalck, Jouan-Rimbaud & Massart, 2000). La principal utilidad de la distancia de Mahalanobis reside en que tiene en cuenta la correlación entre las distintas variables de una observación. Para un vector  $x$  de  $n$  variables, la distancia de Mahalanobis viene dada por

$$d = \sqrt{(x - \mu)^T \Sigma^{-1} (x - \mu)}, \quad (3.2)$$

donde  $\mu$  es el vector de medias de las  $n$  variables y  $\Sigma$  es la matriz de covarianza correspondiente. La distancia de Mahalanobis se puede expresar en términos probabilísticos, convirtiéndose así en una herramienta para detectar *outliers*. Dicha transformación se realiza usando la distribución  $\chi^2$  con  $n$  grados de libertad (Rousseeuw & van Zomeren, 1990). Para determinar si las distancias calculadas son significativas, se hallan los p-valores y se consideran valores anómalos aquellos cuyo p-valor es inferior al umbral dado por un nivel de significancia,  $\alpha$ , establecido de antemano.

En este punto se está en disposición de hallar los valores anómalos de la base de datos relativos a las variables cuantitativas principales con las que se cuenta, el precio y la superficie. Para ello, se agrupa la muestra en doce conjuntos, puesto que se pretende determinar la existencia de observaciones anómalas entre las naves en venta o en alquiler de las comarcas costeras o de interior de cada una de las tres provincias de la Comunidad Valenciana.

En primer lugar, se aplica la técnica univariante del rango intercuartílico sobre la superficie de las naves de la muestra, dando a la constante  $b$  en la ecuación (3.1) un valor de 3. De este modo, únicamente se etiquetan como valores anómalos los extremos. Gráficamente, la detección de *outliers* a partir del rango intercuartílico se basa en el diagrama de caja y bigotes o *box whisker*, que se puede observar en la Figura 3.7.

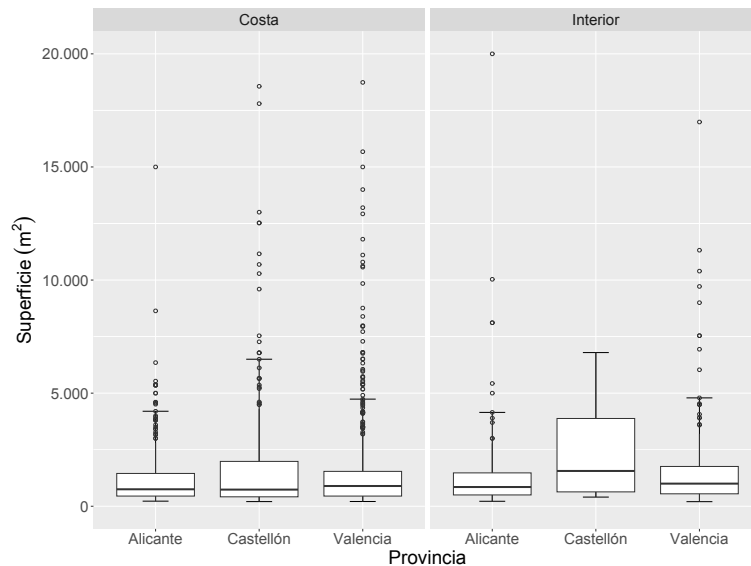


Figura 3.7: Diagrama de caja y bigotes de la variable superficie de las naves en venta

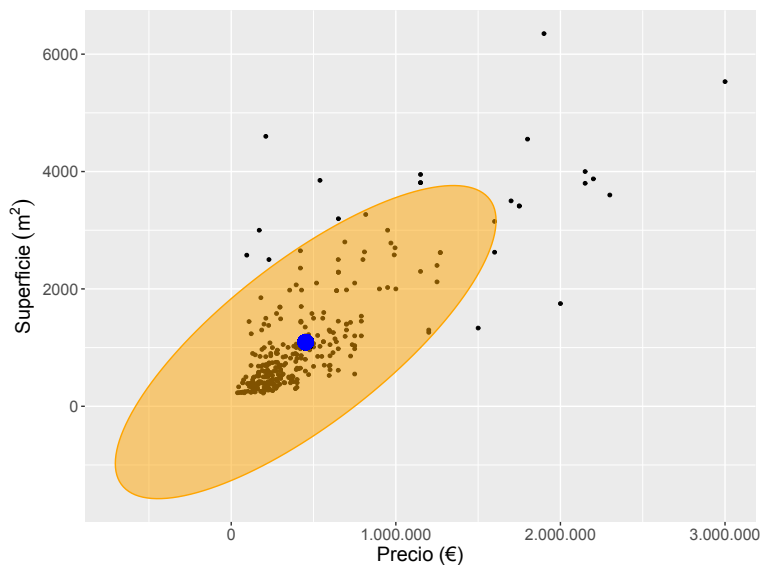
Este diagrama representa la variabilidad de la superficie de las naves en venta. En el mismo, la caja se corresponde con el rango intercuartílico, siendo sus extremos el primer y el tercer cuartil. La línea gruesa dentro de la caja es la mediana de la superficie de las naves en venta, y los bigotes

son los valores extremos del intervalo definido según la [ecuación \(3.1\)](#). Se puede observar, por una parte, que al ser la superficie una variable cuantitativa que no puede tomar valores negativos, el límite inferior es 0. Por otra parte, los puntos que exceden el bigote superior son los valores anómalos.

La [Figura 3.7](#) revela que, si bien el método basado en el rango intercuartílico es útil para suprimir ofertas de inmuebles con una superficie excesivamente elevada, no ocurre lo mismo con inmuebles que presentan una superficie anormalmente pequeña. Por este motivo, se decide eliminar manualmente las naves cuya superficie sea inferior a 200 metros cuadrados, asumiendo que una nave industrial debería disponer de una superficie mayor. Son 113 observaciones las que se eliminan por esta suposición, un 2,3 % de la base de datos original, tras lo que se identifican 119 *outliers* de superficie mediante el rango intercuartílico.

Adicionalmente, para detectar valores anómalos de precio de las naves es conveniente utilizar una técnica multivariante como la distancia de Mahalanobis, ya que dicha variable cuantitativa está correlada con la superficie y esta correlación no puede capturarse mediante enfoques univariante. Aplicando la distancia de Mahalanobis para detectar valores anómalos con un nivel de significancia  $\alpha = 0,025$  (Rousseeuw & van Zomeren, 1990), se identifican 170 ofertas en la muestra.

La distancia de Mahalanobis también puede representarse de forma gráfica como una elipse definida por los parámetros utilizados en el cálculo de dicha distancia. En este sentido, la [Figura 3.8](#) ejemplifica el uso de la distancia de Mahalanobis para hallar los valores anómalos de la fracción de naves industriales a la venta en las comarcas costeras de la provincia de Alicante.



**Figura 3.8:** Distancia de Mahalanobis aplicada a la muestra de naves a la venta en la costa de Alicante

Por un lado, se aprecian las observaciones anómalas, que son aquellas ubicadas fuera de la elipse. Por el otro, puede comprobarse que la mayoría de observaciones se concentran en un rango concreto de precio y superficie. Sin embargo, según este criterio no se pueden descartar observaciones más alejadas, ya que queda demostrado que existe correlación entre ellas.

Aplicados los dos enfoques (univariante y multivariante), se eliminan en total 289 entradas de la base de datos, quedando 2013 naves industriales tras completar el proceso de limpieza de datos. Esto supone un 87 % del total de naves industriales (después de suprimir locales comerciales) y un 40 % de las observaciones iniciales.

### 3.2.3. Estructura final de la base de datos

Del mismo modo que al terminar el proceso de *web scraping*, la [Tabla 3.2](#) resume los atributos que caracterizan la muestra de naves industriales una vez finalizada la limpieza de datos. En comparación con la estructura inicial de la información obtenida (véase la [Tabla 3.1](#)), en la base de datos se reemplazan los campos de descripción y características por las variables generadas durante el proceso de *text mining* descrito en esta sección.

Asimismo, a partir de las variables de precio y superficie de las naves se incorpora una última variable, el precio por metro cuadrado, que se utilizará para determinar la rentabilidad bruta del alquiler según lo explicado en el [Capítulo 2](#). En el caso de las naves en venta, el precio por unidad de superficie es la razón entre estas variables, mientras que para las naves en alquiler el precio por metro cuadrado indica a cuánto asciende la renta anual por unidad de superficie. Por tanto, para calcular el indicador se multiplica previamente por 12 el precio mensual.

**Tabla 3.2:** Resumen de las variables registradas en la base de datos procesada

Variable	Descripción	Tipo	Valores nulos
Operación	Venta o alquiler	Cualitativa	0 %
Precio	Precio de compra (venta) o renta mensual (alquiler)	Cuantitativa	0 %
Superficie	Superficie en metros cuadrados	Cuantitativa	0 %
Precio/m <sup>2</sup>	Precio de compra o renta anual por metro cuadrado	Cuantitativa	0 %
Provincia	Provincia de la nave	Cualitativa	0 %
Comarca	Comarca de la nave	Cualitativa	0 %
Capital	Nave ubicada en una capital de provincia o Elche (sí/no)	Cualitativa	0 %
Consumo	Etiqueta de eficiencia energética asignada a la nave	Cualitativa	87,1 %
Antigüedad	Año de construcción	Cuantitativa	68,1 %
Estado	Estado de la nave (nueva, en buen estado o a reformar)	Cualitativa	0 %
Distribución	Número de estancias (1 = diáfana)	Cuantitativa	23,1 %
Aseos	Aseos disponibles	Cuantitativa	36,7 %
Plantas	Número de plantas	Cuantitativa	70,0 %
Polígono	Nave ubicada en un polígono industrial (sí/no)	Cualitativa	64,3 %
Carretera	Presencia de nudos de comunicación cercanos (sí/no)	Cualitativa	75,9 %
Aparcamiento	Disponibilidad de aparcamiento (sí/no)	Cualitativa	88,0 %

En la [Tabla 3.2](#) también se puede observar que en la base de datos se dispone de variables tanto cualitativas como cuantitativas. Entre las primeras, destacan la variable operación, que indica el mercado al que pertenece una nave (venta o alquiler), y las variables relativas a su ubicación. Entre las variables cuantitativas, las más importantes son las variables continuas de precio y superficie, de ahí que la limpieza de datos se haya basado en esta pareja de atributos.

Por último, se evidencia que muchos atributos contienen un alto porcentaje de valores nulos o faltantes. Mientras que las variables en base a las cuales se han depurado los valores faltantes en las observaciones no presentan valores nulos al haberse garantizado este hecho, la variable de eficiencia energética y las generadas mediante el procesamiento de la descripción de los anuncios están ampliamente incompletas.

La única variable de este grupo consignada en todas las naves es la relativa a su estado. En cuanto al resto, los porcentajes tan altos de valores faltantes se deben a que el anunciante no indica nada sobre estas características. Sin embargo, no es correcto asumir que la ausencia de datos sobre estas variables en una nave implica que la misma no presenta los atributos correspondientes. Así, por ejemplo, del 36,7 % de naves sobre las que se desconoce el número de aseos no se puede concluir que no tengan ninguno, ni aproximar cuántos podrían tener. De la misma forma, que en el 64,3 % de las descripciones de naves no aparezca la palabra *polígono* no significa que algunas de estas naves no puedan estar situadas en polígonos industriales.

Lo único que puede afirmarse con certeza es que el anunciante no ha hecho referencia alguna, siendo este el principal inconveniente del procedimiento de *text mining* llevado a cabo. En consecuencia, el estudio de estas variables con altos porcentajes de valores faltantes se antoja complicado y la probabilidad de obtener resultados interesantes escasa, por lo que no se consideran relevantes como factores explicativos de la ratio de rentabilidad del alquiler de naves.

### **3.3. Análisis exploratorio**

Completado el preprocesamiento de los datos, se dispone de una base de datos que se puede desgranar para obtener una radiografía de la situación del mercado de naves industriales de la Comunidad Valenciana. En esta sección, se analizan las variables relevantes que se han obtenido en las etapas anteriores a través de la estadística descriptiva. La estadística descriptiva, como indica su nombre, es la disciplina que se ocupa de describir de forma cuantitativa un conjunto de datos relativos a una determinada característica, empleando para ello recursos como gráficos, tablas y resúmenes de parámetros estadísticos básicos.

Pese a que hasta el momento se ha hecho referencia a un único conjunto de datos a la hora de preprocesarlos, desde la perspectiva estadística estos datos deben dividirse en dos grupos. Aún habiéndose medido las mismas variables en todas las naves, en realidad se cuenta con dos muestras independientes, una del mercado de venta y otra del mercado de alquiler de esta tipología de inmueble. Por ello, el análisis de las distintas variables se lleva a cabo para las naves en venta y en alquiler por separado, lo que posibilita determinar al mismo tiempo similitudes y diferencias entre las dos muestras.

#### **3.3.1. Distribución geográfica**

Para analizar la localización de las naves, se cuenta con varios niveles de agregación que se corresponden con los distintos niveles de división territorial de la Comunidad Valenciana. Así, el nivel de mayor agregación es el provincial, siendo tres las provincias que forman la Comunidad: Alicante, Castellón y Valencia. A su vez, las provincias se dividen en áreas geográficas más



pequeñas, las comarcas. En la Comunidad Valenciana existen 34 comarcas repartidas por las tres provincias: 9 en Alicante, 8 en Castellón y 17 en Valencia. El nivel de mayor desagregación es el municipal, pero se considera de una granularidad excesiva, de modo que solo se dispone de información en este nivel sobre aquellas naves ubicadas en las capitales de provincia, además de la ciudad de Elche en el caso de la provincia de Alicante. Se considera oportuno incorporar Elche dada su proximidad a la ciudad de Alicante y el similar peso de ambos municipios en términos demográficos y económicos, los factores que precisamente distinguen a una capital de provincia del resto de localidades.

La Figura 3.9 representa la proporción de naves industriales en las distintas provincias de la Comunidad Valenciana, segregada por comarcas. Estos mapas revelan que la oferta de naves no está repartida de manera uniforme. Así, se observa que las naves se localizan principalmente en la costa de la Comunidad y que la oferta en las comarcas del interior es sustancialmente inferior. De hecho, puede comprobarse que la oferta de naves en algunas de estas comarcas es inexistente. En la provincia de Valencia, no constan naves en alquiler en las comarcas del Rincón de Ademuz, el Valle de Cofrentes y Los Serranos. Tampoco hay naves en alquiler en las comarcas de l'Alt Maestrat y Els Ports, pertenecientes a la provincia de Castellón. Además, en otras dos comarcas castellanenses no existe oferta de naves ni en venta ni en alquiler. Se trata de las comarcas del Alto Mijares y El Alcalatén.

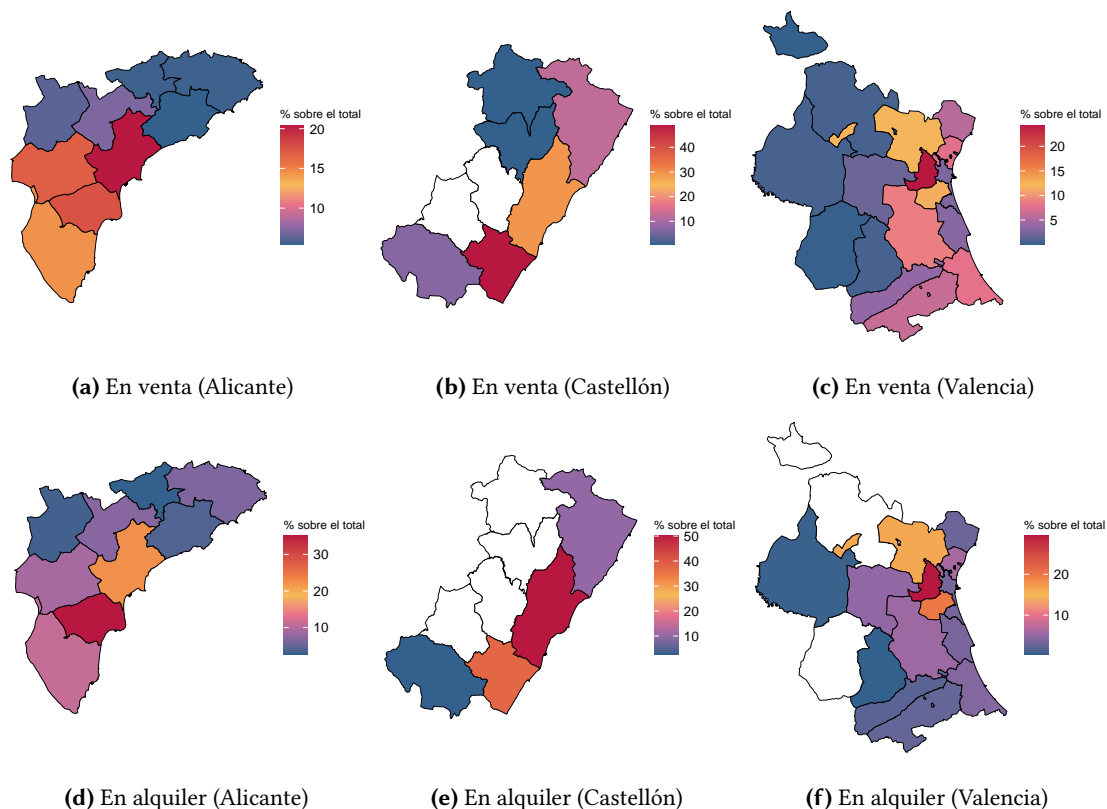


Figura 3.9: Intensidad de la oferta de naves industriales en las provincias de la Comunidad Valenciana

El patrón costa-interior, sin embargo, se ve alterado en Alicante. Estando concentrada la oferta en las comarcas de l'Alacantí y el Baix Vinalopó, en las que se encuentran algunos de los principales municipios alicantinos, una proporción similar de las naves en venta se ubica también en las comarcas colindantes del Vinalopó Mitjà y el Baix Segura, dando lugar a una dualidad norte-sur en esta provincia. Esta tendencia también se observa en el mercado de Alquiler. En cuanto a Castellón, además de la ya citada ausencia de oferta en el interior, las naves se ubican mayoritariamente en las comarcas de la Plana, perteneciendo la ciudad de Castellón a la Plana Alta. En Valencia las naves, sobre todo las ofertadas en alquiler, se localizan principalmente alrededor de la capital, en las comarcas de l'Horta Oest, l'Horta Sud y el Camp del Túria. La [Tabla 3.3](#) resume la distribución de la muestra de naves a lo largo de la geografía valenciana, cuantificando la oferta de naves en venta y en alquiler en cada comarca respecto al total de cada provincia.

**Tabla 3.3:** Oferta de naves industriales por provincia en la Comunidad Valenciana

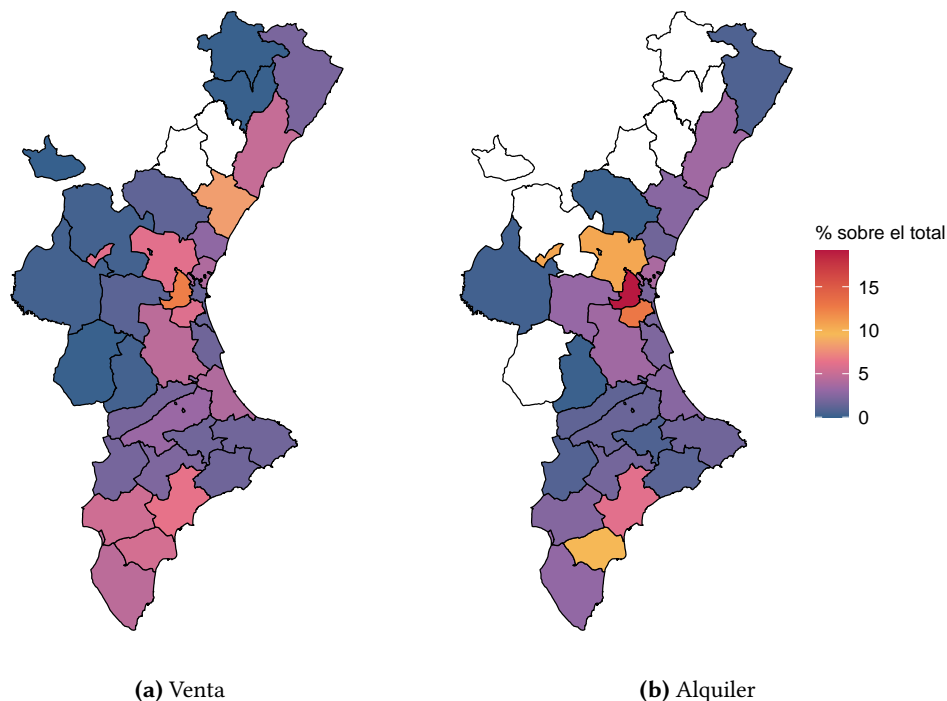
	Naves en venta		Naves en alquiler	
	Recuento	% sobre provincia	Recuento	% sobre provincia
<b>Alicante</b>				
L'Alt Vinalopó	29	6,37 %	5	3,18 %
El Baix Segura	66	14,51 %	17	10,83 %
El Baix Vinalopó	80	17,58 %	55	35,03 %
El Comtat	26	5,71 %	4	2,55 %
L'Alacantí	93	20,44 %	35	22,29 %
L'Alcoià	33	7,25 %	11	7,01 %
La Marina Alta	26	5,71 %	10	6,37 %
La Marina Baixa	25	5,49 %	6	3,82 %
El Vinalopó Mitjà	77	16,92 %	14	8,92 %
<b>Total</b>	<b>455</b>	<b>100 %</b>	<b>157</b>	<b>100 %</b>
<b>Castellón</b>				
L'Alt Maestrat	2	0,82 %	-	-
El Alto Palancia	18	7,35 %	1	2,50 %
El Baix Maestrat	31	12,65 %	4	10,00 %
Els Ports	2	0,82 %	-	-
La Plana Alta	72	29,39 %	20	50,00 %
La Plana Baixa	120	48,98 %	15	37,50 %
<b>Total</b>	<b>245</b>	<b>100 %</b>	<b>40</b>	<b>100 %</b>
<b>Valencia</b>				
El Camp de Morvedre	42	5,64 %	10	2,70 %
El Camp del Túria	89	11,95 %	60	16,17 %
El Rincón de Ademuz	1	0,13 %	-	-
El Valle de Cofrentes-Ayora	2	0,27 %	-	-
La Hoya de Buñol	16	2,15 %	18	4,85 %
L'Horta Nord	59	7,92 %	22	5,93 %
L'Horta Oest	181	24,30 %	109	29,38 %
L'Horta Sud	86	11,54 %	73	19,68 %
La Canal de Navarrés	6	0,81 %	1	0,27 %
La Costera	30	4,03 %	7	1,89 %
La Plana de Utiel-Requena	6	0,81 %	2	0,54 %
La Vall d'Albaida	48	6,44 %	9	2,43 %
Los Serranos	6	0,81 %	-	-
La Ribera Alta	66	8,86 %	20	5,39 %
La Ribera Baixa	25	3,36 %	13	3,50 %
La Safor	60	8,05 %	15	4,04 %
Valencia	22	2,95 %	12	3,23 %
<b>Total</b>	<b>745</b>	<b>100 %</b>	<b>371</b>	<b>100 %</b>

Del estudio se evidencia que la oferta de venta en la Comunidad Valenciana es ampliamente superior a la del alquiler. En el conjunto de la Comunidad, un 71,8 % de las naves pertenece a la primera categoría y solo un 28,2 % a la segunda. La mayor diferencia entre mercados se da en la provincia de Castellón, con tan solo un 13,6 % de las naves en alquiler, mientras que la menor diferencia ocurre en Valencia, con un 66,8 % y 33,2 %, respectivamente.

Asimismo, se refrendan las observaciones anteriores sobre la [Figura 3.9](#). En la provincia de Alicante, las cuatro comarcas del cono sur concentran el 69,5 % de las naves en venta y el 77,1 % de las naves en alquiler. Mientras que l'Alacantí y el Baix Segura presentan un equilibrio entre venta y alquiler en términos relativos, en el Baix Vinalopó la oferta está orientada hacia el alquiler y en el Vinalopó Mitjà, hacia la venta.

En Castellón, las comarcas de la Plana absorben el 80 % de la oferta de venta y el 90 % de la del alquiler, aproximadamente. En términos absolutos, la oferta de naves en alquiler es muy reducida, con solo 40 inmuebles en toda la provincia. Igualmente, en las comarcas sin oferta de alquiler (l'Alt Maestrat y Els Ports) solo hay 4 naves en venta. En cuanto a Valencia, se observa que l'Horta Oest concentra alrededor de una cuarta parte de la oferta de venta y cerca de una tercera parte del alquiler. Asimismo, el Rincón de Ademuz y el Valle de Cofrentes, sin naves en alquiler, solo cuentan con una y dos naves en venta, respectivamente.

La [Figura 3.10](#) ofrece una fotografía de la distribución geográfica de las naves industriales que abarca toda la Comunidad Valenciana, completada por la [Tabla 3.4](#), que ofrece datos sobre la intensidad de la oferta de naves por comarca respecto al total de la Comunidad Valenciana.



**Figura 3.10:** Intensidad de la oferta de naves industriales en la Comunidad Valenciana por comarca

En línea con los comentarios anteriores, se evidencia que la oferta de naves en la Comunidad Valenciana se concentra en dos zonas: las comarcas vecinas de Valencia y las cuatro comarcas del sur de Alicante, en las que se localizan las ciudades de Alicante y Elche. La comarca valenciana de l’Horta Oest presenta el mayor número de naves de la Comunidad, un 12,5 % y un 19,2 % de las ofertadas en venta y en alquiler, respectivamente. Junto al Camp del Túria y l’Horta Sud, elevan el número de naves disponibles hasta situar Valencia como la provincia con mayor proporción de estos inmuebles. Las comarcas de la Plana de Castellón, que presentan la intensidad de oferta más elevada de la provincia, apenas destacan en el mercado de alquiler de la Comunidad, siendo su peso en términos relativos en este segmento equiparable al de algunas comarcas de interior.

**Tabla 3.4:** Distribución por comarcas de la oferta de naves industriales de la Comunidad Valenciana

	Naves en venta		Naves en alquiler	
	Recuento	% sobre el total	Recuento	% sobre el total
<b>Alicante</b>				
L’Alt Vinalopó	29	2,01 %	5	0,88 %
El Baix Segura	66	4,57 %	17	2,99 %
El Baix Vinalopó	80	5,54 %	55	9,68 %
El Comtat	26	1,80 %	4	0,70 %
L’Alacantí	93	6,44 %	35	6,16 %
L’Alcoià	33	2,28 %	11	1,94 %
La Marina Alta	26	1,80 %	10	1,76 %
La Marina Baixa	25	1,73 %	6	1,06 %
El Vinalopó Mitjà	77	5,33 %	14	2,46 %
<b>Total</b>	<b>455</b>	<b>31,50 %</b>	<b>157</b>	<b>27,63 %</b>
<b>Castellón</b>				
L’Alt Maestrat	2	0,14 %	-	-
El Alto Palancia	18	1,25 %	1	0,18 %
El Baix Maestrat	31	2,15 %	4	0,70 %
Els Ports	2	0,14 %	-	-
La Plana Alta	72	4,98 %	20	3,52 %
La Plana Baixa	120	8,30 %	15	2,64 %
<b>Total</b>	<b>245</b>	<b>16,96 %</b>	<b>40</b>	<b>7,04 %</b>
<b>Valencia</b>				
El Camp de Morvedre	42	2,91 %	10	1,76 %
El Camp del Túria	89	6,16 %	60	10,56 %
El Rincón de Ademuz	1	0,07 %	-	-
El Valle de Cofrentes-Ayora	2	0,14 %	-	-
La Hoya de Buñol	16	1,11 %	18	3,17 %
L’Horta Nord	59	4,08 %	22	3,87 %
L’Horta Oest	181	12,53 %	109	19,19 %
L’Horta Sud	86	5,95 %	73	12,85 %
La Canal de Navarrés	6	0,42 %	1	0,18 %
La Costera	30	2,08 %	7	1,23 %
La Plana de Utiel-Requena	6	0,42 %	2	0,35 %
La Vall d’Albaida	48	3,32 %	9	1,58 %
Los Serranos	6	0,42 %	-	-
La Ribera Alta	66	4,57 %	20	3,52 %
La Ribera Baixa	25	1,73 %	13	2,29 %
La Safor	60	4,15 %	15	2,64 %
Valencia	22	1,52 %	12	2,11 %
<b>Total</b>	<b>745</b>	<b>51,58 %</b>	<b>371</b>	<b>65,30 %</b>
<b>Total Comunidad Valenciana</b>				
	<b>1445</b>	<b>100 %</b>	<b>568</b>	<b>100 %</b>

Por último, se analiza la oferta de naves industriales en las capitales de provincia y la ciudad de Elche. 240 naves, 146 en venta y 94 en alquiler, se ubican en estos términos municipales, suponiendo el 10,1 % y el 16,5 % de la oferta de venta y alquiler, respectivamente. La [Tabla 3.5](#) resume las naves en venta y en alquiler disponibles en cada núcleo urbano. Se observa que el área Alicante/Elche es la que más naves presenta. Además, concentra una parte reseñable de la oferta, tanto de venta como de alquiler, de las comarcas de l'Alacantí y el Baix Vinalopó, al igual que en el conjunto de la provincia. No obstante, estos datos deben analizarse teniendo en cuenta que se trata de la unión de la oferta de dos grandes ciudades. Pese a todo, remarca la relevancia del núcleo Alicante/Elche en el mercado de naves industriales de esta provincia.

**Tabla 3.5:** Oferta de naves industriales en las capitales de provincia de la Comunidad Valenciana

	Naves en venta			Naves en alquiler		
	Recuento	% sobre provincia	% sobre comarca	Recuento	% sobre provincia	% sobre comarca
Alicante/Elche	104	22,86 %	60,12 %	75	47,77 %	83,33 %
Castellón de la Plana	23	9,39 %	31,94 %	9	22,50 %	45,00 %
Valencia	19	2,55 %	86,36 %	10	2,70 %	83,33 %

Mientras tanto, en la ciudad de Castellón se comprueba que las naves situadas en su término municipal no son mayoritarias en la comarca a la que pertenece, y mucho menos a nivel provincial. Al analizar Valencia, el altísimo peso relativo de la oferta de naves en la ciudad respecto a la comarca se explica por la prácticamente total coincidencia geográfica entre ambas divisiones territoriales. En cambio, la intensidad de la oferta es testimonial respecto al conjunto de la provincia, puesto que las naves se concentran en la zona metropolitana formada por las comarcas colindantes enumeradas anteriormente.

### 3.3.2. Precio

El precio es una de las variables cuantitativas disponibles en las muestras. En el caso de las naves en venta, el precio representa la cantidad de dinero a pagar para adquirirlas en propiedad, mientras que en la contraparte de alquiler, la variable precio indica la renta mensual a abonar para arrendarlas.

La [Tabla 3.6](#) presenta los estadísticos que resumen la distribución del precio de las naves en venta. Análogamente, la [Tabla 3.7](#) replica los mismos parámetros derivados de la muestra de naves en alquiler. El rango de precios es amplio, existiendo una oferta de compra por 15.000 euros, mientras que la nave más cara asciende a 2,5 millones de euros. Del mismo modo ocurre con la oferta de alquiler, encontrándose un anuncio de una nave en alquiler por 100 euros mensuales, mientras que el arrendamiento más oneroso supone un desembolso de 118.504 euros al mes.

En cuanto al precio medio, este se sitúa en 321.978 euros en el mercado de venta, mientras que alcanza los 6.147 euros mensuales en el de alquiler. Al mismo tiempo, el 50 % de las naves en venta tiene un precio entre 155.000 y 395.000 euros. Respecto a las naves en alquiler, la mitad de los inmuebles ofertados puede arrendarse por un precio entre 1.200 y 4.500 euros mensuales.

Por otra parte, del coeficiente de variación se desprende la mayor dispersión de los precios en el mercado de alquiler que entre la oferta de naves en venta. En cualquier caso, los elevados

coeficientes de variación (86 % y 212 %, respectivamente) denotan que la media de los precios no es representativa del conjunto. Respecto a la simetría de las distribuciones, el precio de las naves en alquiler presenta un sesgo mucho mayor que el de las naves en venta, como se desprende de los respectivos coeficientes de simetría. En cuanto a la curtosis, aunque la distribución es leptocúrtica o puntiaguda en ambos casos (coeficiente mayor que 3), la del alquiler es superior, en línea con los indicadores anteriores.

**Tabla 3.6:** Resumen de la variable precio de las naves en venta

<b>Mínimo</b>	15000 €
<b>Percentil 25</b>	155000 €
<b>Media</b>	321977,94 €
<b>Mediana</b>	239000 €
<b>Percentil 75</b>	395000 €
<b>Máximo</b>	2500000 €
<b>Desviación típica</b>	275928,53 €
<b>Coeficiente de variación</b>	0,857
<b>Coeficiente de simetría</b>	2,694
<b>Coeficiente de curtosis</b>	10,613

**Tabla 3.7:** Resumen de la variable precio de las naves en alquiler

<b>Mínimo</b>	100 €
<b>Percentil 25</b>	1200 €
<b>Media</b>	6147,25 €
<b>Mediana</b>	2000 €
<b>Percentil 75</b>	4500 €
<b>Máximo</b>	118504 €
<b>Desviación típica</b>	13007,04 €
<b>Coeficiente de variación</b>	2,116
<b>Coeficiente de simetría</b>	5,047
<b>Coeficiente de curtosis</b>	31,737

La [Figura 3.11](#) complementa gráficamente el análisis anterior. Tal y como era de esperar, se comprueba que el precio es una variable sesgada de por sí, en tanto que no puede tomar valores negativos. La curva de densidad del precio de las naves en venta y en alquiler revela, por un lado, la asimetría positiva, situándose el precio de la mayoría de los inmuebles ofertados en la parte baja del rango definido previamente. Por otro lado, se aprecia el gran apuntamiento, concentrándose los precios en torno a los valores centrales de la distribución. Estas características se aprecian de forma más marcada en el precio de la oferta de naves en alquiler. También se puede observar la distancia entre la media y la mediana.

La evidente asimetría positiva de la variable precio explica que en la muestra aparezcan extrañas ofertas en el contexto de naves industriales, como la correspondiente al mínimo precio de alquiler. Se trata de un pequeño almacén industrial para reformar en Castellón y se puede arrendar por 100 euros mensuales. Dado que los valores centrales son muy próximos a 0, a través de los métodos empleados para suprimir valores anómalos no se consigue eliminar tales observaciones.

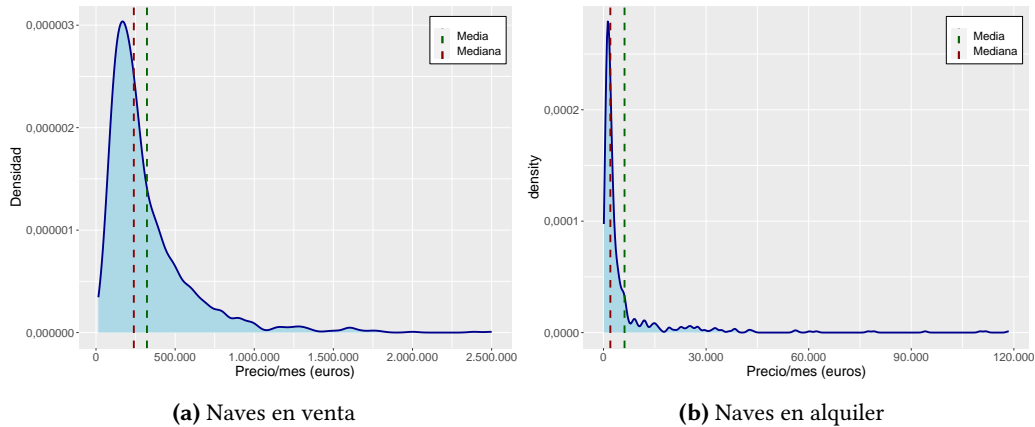


Figura 3.11: Curva de densidad de la variable precio

Analizada la distribución del precio de la oferta de venta y alquiler de naves industriales en su totalidad, a continuación se estudia geográficamente. La Figura 3.12 presenta el precio medio de las naves en venta y en alquiler en cada comarca de la Comunidad Valenciana. Mientras que se observa una heterogeneidad en el precio medio de la oferta de venta, habiendo comarcas donde se supera el medio millón de euros, y ningún patrón aparente, la situación en el mercado de alquiler es la contraria. En gran parte de la Comunidad Valenciana, el precio medio de la oferta de alquiler de naves es similar y se sitúa en la parte baja del rango, como ya dejaba entrever la Figura 3.11 (b).

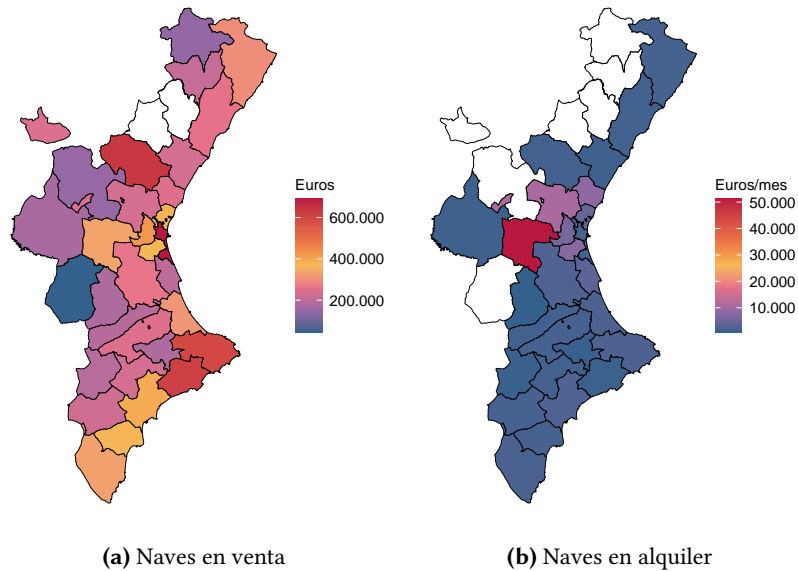


Figura 3.12: Precio medio de las naves industriales por comarca

Sin embargo, en algunas comarcas de Valencia (l'Horta, el Camp de Morvedre y el Camp del Túria), el precio medio del alquiler de naves es notablemente superior al del resto de la Comunidad, situándose entre los 5.700 y los 10.700 euros mensuales, aproximadamente. Por encima de estas

se sitúa, en cuanto al precio medio del alquiler, la Hoya de Buñol, con un precio medio mensual alrededor de los 50.000 euros. Esta importante diferencia podría deberse a la escasa oferta de naves de alquiler en las comarcas de interior que se ha explicado en el apartado anterior.

Por último, la [Tabla 3.8](#) detalla la distribución del precio de las naves en venta y en alquiler por comarca y provincia. Del lado de la venta destaca la similitud en los precios medios y medianos de las provincias de Alicante y Valencia, mientras que entre la oferta de alquiler cabe subrayar nuevamente los elevados precios medios en la provincia de Valencia, impulsados por los datos en la comarca de la Hoya de Buñol, el Camp del Túria, el Camp de Morvedre y, en menor medida, l’Horta Sud. En cambio, el precio mediano está más contenido. Del lado del alquiler también se observa la similitud en el precio mediano de las naves en Alicante y Castellón.

**Tabla 3.8:** Precio de la oferta de naves industriales por comarca

	Precio de venta (€)		Renta mensual (€)	
	Medio	Mediano	Medio	Mediano
<b>Alicante</b>				
L’Alt Vinalopó	192.537	125.000	1.015	1.000
El Baix Segura	327.324	250.000	2.205	2.000
El Baix Vinalopó	374.949	277.600	1.985	1.700
El Comtat	180.874	157.500	972	1.000
L’Alacantí	391.837	315.000	2.539	2.500
L’Alcoià	232.377	166.300	1.573	1.500
La Marina Alta	591.377	488.200	2.338	1.575
La Marina Baixa	609.968	420.000	1.033	1.050
El Vinalopó Mitjà	224.036	190.000	1.889	1.500
	<b>338.177</b>	<b>250.000</b>	<b>2.024</b>	<b>1.600</b>
<b>Castellón</b>				
L’Alt Maestrat	210.000	210.000	-	-
El Alto Palancia	631.222	590.000	1.500	1.500
El Baix Maestrat	297.015	195.000	1.838	1.475
Els Ports	146.000	146.000	-	-
La Plana Alta	258.286	165.000	1.554	1.500
La Plana Baixa	227.716	176.350	1.373	1.200
	<b>274.302</b>	<b>184.400</b>	<b>1.513</b>	<b>1.475</b>
<b>Valencia</b>				
El Camp de Morvedre	246.547	190.002	8.063	1.875
El Camp del Túria	230.034	187.000	10.898	4.500
El Rincón de Ademuz	241.500	241.500	-	-
El Valle de Cofrentes-Ayora	42.825	42.825	-	-
La Hoya de Buñol	331.301	257.100	51.136	56.166
L’Horta Nord	378.568	290.000	4.310	1.850
L’Horta Oest	424.140	350.000	5.715	3.000
L’Horta Sud	372.366	279.950	7.599	2.900
La Canal de Navarrés	178.847	174.240	600	600
La Costera	191.056	157.500	1.586	1.200
La Plana de Utiel-Requena	171.829	177.600	1.250	1.250
La Vall d’Albaida	247.565	180.000	1.931	1.496
Los Serranos	151.233	95.700	-	-
La Ribera Alta	261.977	209.500	2.445	850
La Ribera Baixa	199.153	189.000	3.620	2.000
La Safor	308.956	247.050	2.313	1.600
Valencia	689.375	482.500	1.968	1.595
	<b>327.763</b>	<b>245.000</b>	<b>8.392</b>	<b>2.500</b>



En la provincia de Alicante, el precio medio de venta más alto se da en la Marina, mientras que El Comtat presenta el más bajo. Sin embargo, si se presta atención a los precios medianos, se observa que los más altos corresponden a las comarcas con más oferta en dicha provincia junto a las anteriores. En el mercado de alquiler, los precios son similares en todas las comarcas, a excepción de l'Alt Vinalopó, El Comtat y la Marina Baixa, que tienen los precios más bajos. Por último, en la provincia de Castellón resulta llamativo el alto precio de venta en la comarca del Alto Palancia, aunque de nuevo este hecho podría estar motivado por la escasa oferta en la zona.

### 3.3.3. Superficie

La superficie es otra de las variables cuantitativas mediante la que se puede describir la naturaleza de las naves industriales. Procediendo de la misma forma que con el precio, el análisis de la distribución de esta variable comienza con el cálculo de los estadísticos resumen. Los correspondientes a la superficie de las naves en venta se encuentran en la [Tabla 3.9](#), mientras que la [Tabla 3.10](#) recoge los relativos a la superficie de las naves en alquiler.

**Tabla 3.9:** Resumen de la variable superficie de las naves en venta

<b>Mínimo</b>	203 m2
<b>Percentil 25</b>	435 m2
<b>Media</b>	995,5 m2
<b>Mediana</b>	723 m2
<b>Percentil 75</b>	1264 m2
<b>Máximo</b>	7537 m2
<b>Desviación típica</b>	809,51 m2
<b>Coefficiente de variación</b>	0,813
<b>Coefficiente de simetría</b>	2,316
<b>Coefficiente de curtosis</b>	8,761

En su conjunto, la distribución de la superficie de las naves presenta unas características similares a las del precio. El rango de superficies observado es amplio, sobre todo el correspondiente a las naves en alquiler. En la parte baja se comprueba que, tal y como se había determinado durante la depuración de la base de datos original, no existen inmuebles con una superficie inferior a 200 metros cuadrados, siendo la superficie mínima de 203 metros cuadrados en ambos casos.

**Tabla 3.10:** Resumen de la variable superficie de las naves en alquiler

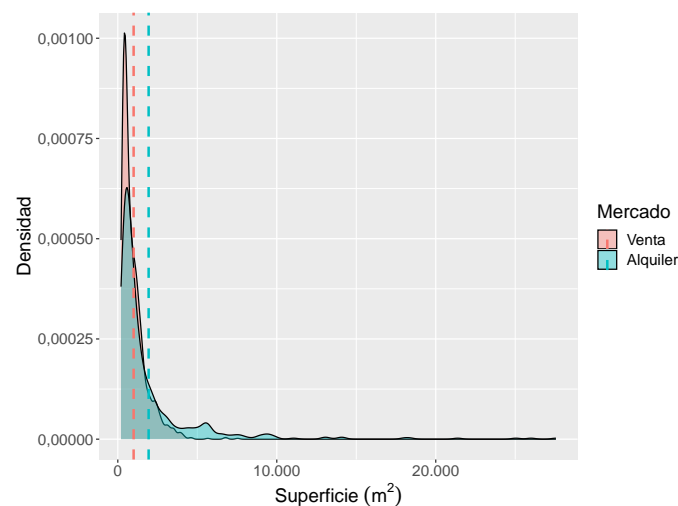
<b>Mínimo</b>	203 m2
<b>Percentil 25</b>	516 m2
<b>Media</b>	1938,68 m2
<b>Mediana</b>	900 m2
<b>Percentil 75</b>	2000 m2
<b>Máximo</b>	27559 m2
<b>Desviación típica</b>	3039,25 m2
<b>Coefficiente de variación</b>	1,568
<b>Coefficiente de simetría</b>	4,538
<b>Coefficiente de curtosis</b>	27,188

En cuanto a la superficie máxima, se observa que en el caso de las naves en venta asciende a 7.537 metros cuadrados, mientras que la nave en alquiler más grande tiene 27.559 metros cuadrados. La superficie media de las naves en venta es de 995 metros cuadrados; en cambio, se sitúa sobre los 1.940 metros cuadrados para las naves en alquiler. Sin embargo, existe una mayor similitud entre las medianas de ambos conjuntos, 723 y 900 metros cuadrados, respectivamente. Por otra parte, el 50 % de las naves en venta tiene una superficie comprendida entre 435 y 1.265 metros cuadrados, mientras que el 50 % de la oferta de naves en alquiler cuenta con entre 516 y 2.000 metros cuadrados.

Profundizando en el análisis de la distribución de la superficie de las naves industriales en su conjunto, se pueden destacar los mismos factores que en el caso del precio en cuanto a la dispersión o la simetría. Así pues, la superficie también presenta una elevada dispersión respecto a la media, siendo más notoria en el segmento de alquiler. Por otra parte, se analiza nuevamente una variable sesgada que no puede tomar valores negativos distribuida con una marcada asimetría positiva (la media es mayor que la mediana), ya que la superficie de la mayor parte de las naves se sitúa en la parte baja de los rangos definidos anteriormente.

En este sentido, el sesgo positivo en la superficie de la oferta de alquiler es mayor, con un coeficiente de simetría de 4,5 que es superior al registrado para la superficie de las naves en venta, 2,3. Este es el motivo por el cual ha sido necesario establecer manualmente la superficie mínima a considerar. Si no, habría resultado imposible suprimir estas ofertas con las técnicas de detección de valores anómalos utilizadas, al no desviarse lo suficiente de la parte central del conjunto.

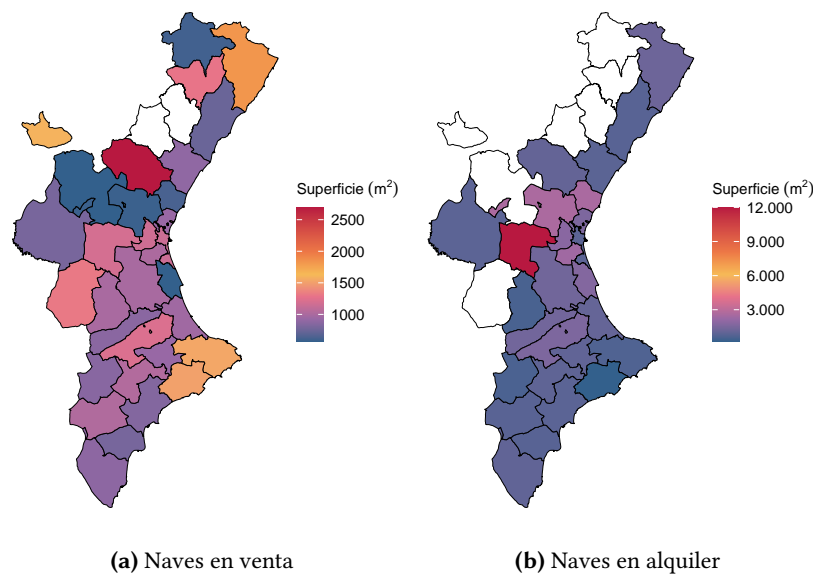
Adicionalmente, la [Figura 3.13](#) contiene la distribución de la superficie de las naves en venta y en alquiler. Se observa la asimetría de la superficie de las dos muestras y cómo la dispersión respecto a la media de la superficie entre las naves en alquiler es superior a la que se da entre las naves en venta, hecho que también se deriva de los coeficientes de variación que aparecen en la [Tabla 3.9](#) y la [Tabla 3.10](#), respectivamente.



**Figura 3.13:** Curvas de densidad de la variable superficie en el mercado de oferta y de alquiler

Una vez estudiada la distribución de la superficie de las naves industriales en venta y en alquiler en el conjunto de la Comunidad Valenciana, a continuación se pretende identificar patrones acotados geográficamente. La [Figura 3.14](#) representa la superficie media de la oferta de naves en venta y en alquiler por comarcas. En cuanto a las naves en venta, la superficie media más alta se da en la comarca castellanense del Alto Palancia con aproximadamente 2.700 metros cuadrados, seguida por el también castellanense Baix Maestrat (1.834), el valenciano Rincón de Ademuz (1.600) y las comarcas alicantinas de la Marina, con cerca de 1.500 metros cuadrados. No obstante, debe tenerse en cuenta que, la superficie media en el Rincón de Ademuz corresponde a la única nave ofertada en dicha comarca. En cambio, las superficies medias más reducidas se hallan en las comarcas del norte de Valencia, la Ribera Baixa y Els Ports, esta última en Castellón.

Si se presta atención a la superficie de las naves en alquiler, se observa que la comarca valenciana de la Hoya de Buñol presenta la media más alta, por encima de los 12.000 metros cuadrados, muy diferenciada de la superficie media en alquiler en el resto de zonas, pero también de las medias en el segmento de venta. Al mismo tiempo, las superficies medias más altas en este mercado corresponden a las comarcas de Valencia que presentaban los valores más bajos entre la oferta de naves en venta. También cabe destacar el valor mínimo y alejado del resto que se registra en la Marina Baixa (Alicante), 262 metros cuadrados. Pese a todo, este dato debe contextualizarse porque en dicha comarca únicamente constan seis naves en alquiler (véase la [Tabla 3.4](#)).



**Figura 3.14:** Superficie media de las naves industriales por comarca

Los datos anteriores, así como los correspondientes al resto de comarcas, pueden consultarse en la [Tabla 3.11](#). En definitiva, observando los valores recogidos en la misma se puede comprobar que los lugares en los que se identifican los precios más altos en términos medios y medianos (véase la [Tabla 3.8](#)) son mayoritariamente los mismos que presentan las naves de mayor superficie. Este hecho se aprecia, sobre todo, en el mercado de alquiler, en las comarcas valencianas comentadas anteriormente. En el mercado de venta, esta fuerte relación se da en las comarcas alicantinas de la Marina, así como en las castellanenses del Alto Palancia y el Baix Maestrat. Además, cabe resaltar

que en estas comarcas con las superficies medias más altas es donde mayor diferencia existe entre la superficie de las naves en venta y en alquiler en términos medios.

**Tabla 3.11:** Superficie de la oferta de naves industriales por comarca

	Superficie de venta (m <sup>2</sup> )		Superficie de alquiler (m <sup>2</sup> )	
	Media	Mediana	Media	Mediana
<b>Alicante</b>				
L'Alt Vinalopó	873	530	632	521
El Baix Segura	890	713	1.052	720
El Baix Vinalopó	803	627	860	650
El Comtat	936	694	1.052	1.005
L'Alacantí	851	620	937	809
L'Alcoià	1.046	843	992	635
La Marina Alta	1.544	1.025	714	588
La Marina Baixa	1.515	1.100	262	255
El Vinalopó Mitjà	1.037	850	879	733
	<b>976</b>	<b>708</b>	<b>874</b>	<b>650</b>
<b>Castellón</b>				
L'Alt Maestrat	1.300	1.300	-	-
El Alto Palancia	2.696	2.134	1.100	1.100
El Baix Maestrat	1.834	1.233	1.294	1.376
Els Ports	621	621	-	-
La Plana Alta	757	498	854	575
La Plana Baixa	905	607	920	614
	<b>1.112</b>	<b>653</b>	<b>929</b>	<b>685</b>
<b>Valencia</b>				
El Camp de Morvedre	641	494	2.630	1.270
El Camp del Túria	600	432	2.651	1.510
El Rincón de Ademuz	1.600	1.600	-	-
El Valle de Cofrentes-Ayora	1.324	1.324	-	-
La Hoya de Buñol	1.194	1.200	12.009	13.062
L'Horta Nord	943	887	1.651	632
L'Horta Oest	1.168	1.066	1.895	1.050
L'Horta Sud	1.018	800	2.420	1.300
La Canal de Navarrés	1.004	1.076	600	600
La Costera	881	581	1.399	870
La Plana de Utiel-Requena	811	710	950	950
La Vall d'Albaida	1.230	1.000	1.639	1.350
Los Serranos	582	488	-	-
La Ribera Alta	1.000	698	1.377	525
La Ribera Baixa	581	520	1.753	1.625
La Safor	970	994	1.115	900
Valencia	1.154	886	957	940
	<b>969</b>	<b>787</b>	<b>2.498</b>	<b>1.070</b>

En líneas generales, los datos de la [Tabla 3.11](#) reflejan igualmente la asimetría en la distribución de la superficie particularizando por provincias y comarcas, dado que en prácticamente todos los casos el valor medio es superior a la mediana. La superficie media y, sobre todo, la superficie mediana de las naves en venta son relativamente similares en las tres provincias de la Comunidad, aunque destaca la diferencia entre estos dos valores en el caso de Castellón. Por lo que respecta al segmento del alquiler, Alicante y Castellón presentan unos valores medios y medianos parecidos entre sí y que no distan en exceso de los que se dan en el mercado de venta. Sin embargo, la diferencia es amplia en comparación con Valencia debido a las grandes naves ubicadas en las comarcas de esta provincia que se han enumerado con anterioridad.

Para concluir con el análisis de la superficie de la oferta de naves industriales de la Comunidad Valenciana, conviene poder afirmar con mayor rigurosidad si existen diferencias entre los distintos valores medios y medianos calculados. Es especialmente importante realizar esta comparación entre las ofertas de naves en venta y alquiler de una zona geográfica concreta, dado que la ratio de rentabilidad que se pretende calcular en última instancia será más fiable cuanto más similares sean las naves en términos de superficie. En otras palabras, el precio por metro cuadrado de cada mercado que se utiliza en el cálculo de la rentabilidad vendrá dado por naves del mismo tamaño.

Por ello, cabe aplicar alguna técnica para descubrir si dichas diferencias son estadísticamente significativas. En este sentido, la estadística inferencial aporta los métodos para poder alcanzar tales conclusiones, ya que como indica su nombre, permite determinar las propiedades de una población (en este caso, el mercado de naves industriales) a partir de las muestras. En concreto, se recurre a un contraste de hipótesis para esclarecer si la distribución de la superficie es la misma en los dos mercados.

Puesto que esta variable no sigue una distribución normal y tiene una marcada asimetría, como se puede observar en la [Figura 3.13](#), se descarta el empleo de pruebas paramétricas porque estas asumen la normalidad de los datos (Butler, 1985). Así pues, una prueba como la *t* de Student para estudiar si las medias de la superficie de las naves en venta y en alquiler son significativamente distintas no debería usarse porque es sensible a la normalidad. En cambio, se puede utilizar una prueba no paramétrica, ya que aceptan una distribución libre. El equivalente no paramétrico a la prueba anterior es la prueba U de Mann-Whitney y el estadístico analizado en este caso es la mediana.

Por tanto, para realizar la prueba U de Mann-Whitney se plantean las siguientes hipótesis:

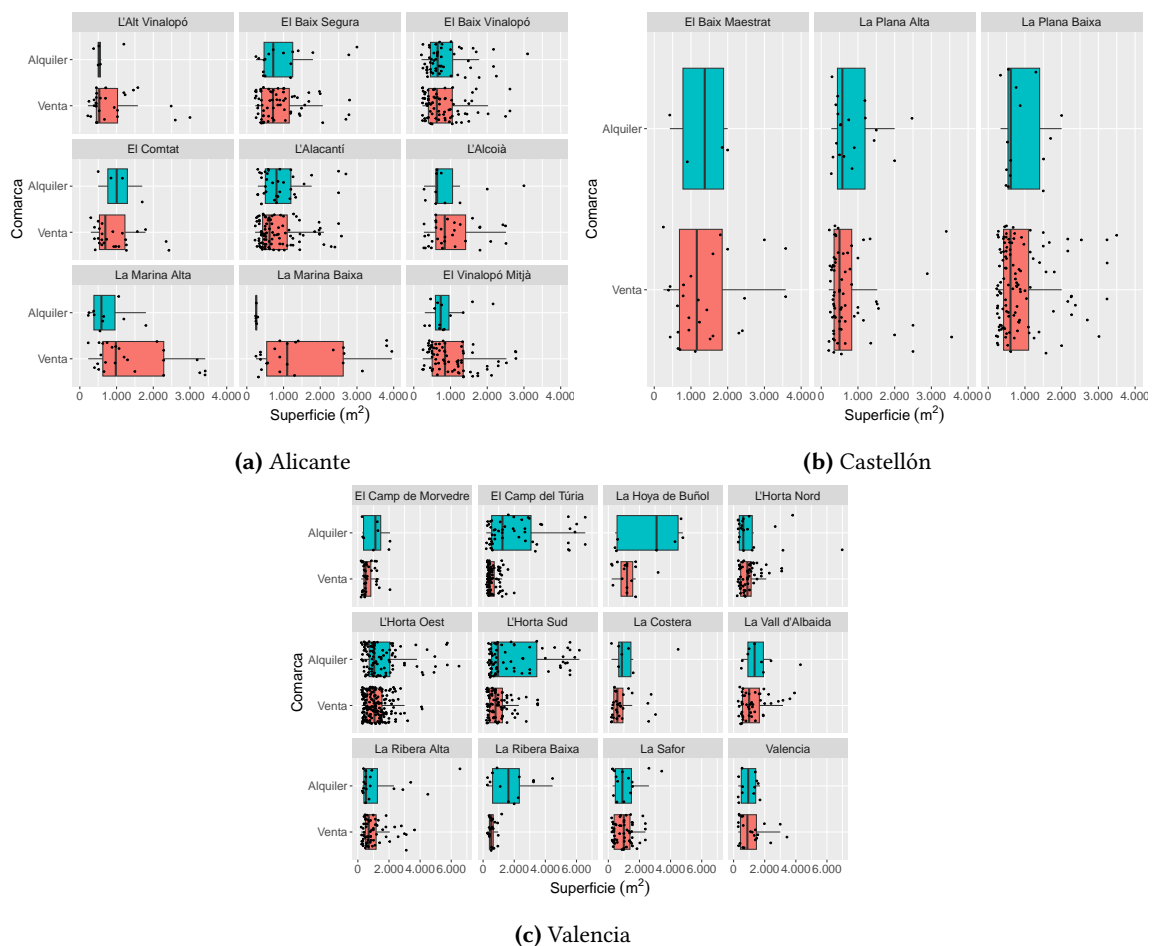
$$\begin{aligned} H_0 : P(X > Y) &= \frac{1}{2} \\ H_1 : P(X > Y) &\neq \frac{1}{2} \end{aligned} \tag{3.3}$$

Siendo *X* e *Y* las observaciones de superficie de las muestras de naves en venta y en alquiler, respectivamente, y ordenadas ascendentemente, la hipótesis nula,  $H_0$ , se traduce en que cada observación de una muestra tiene un 50 % de probabilidad de ser superior que la observación correspondiente de la otra muestra. Esto es, la probabilidad de que una observación sea mayor que la análoga es la misma que la probabilidad del suceso opuesto. Por el contrario, la hipótesis alternativa,  $H_1$ , plantea que la probabilidad es superior o inferior a dicho 50 %. Si no existe una diferencia significativa, es decir, no se puede rechazar la hipótesis nula, es de esperar que las medianas sean similares y que las diferencias se deban a la aleatoriedad.

La prueba U de Mann-Whitney se aplica en aquellas comarcas que disponen de suficiente oferta de naves industriales en venta y en alquiler para establecer si existen diferencias estadísticamente significativas en términos de superficie de los inmuebles entre los dos mercados. Se escoge un nivel de significancia  $\alpha = 0,05$  con el que se comparan los p-valores resultantes tras aplicar la prueba. En las comarcas alicantinas de la Marina Baixa y la Marina Alta, así como en las comarcas valencianas del Camp de Morvedre, Camp del Túria, la Hoya de Buñol, l'Horta Oest, l'Horta Sud y la Ribera Baixa, el p-valor obtenido es inferior al 5 %, de modo que debe rechazarse la

hipótesis nula y admitir que la superficie de las naves en venta y en alquiler en estas localizaciones difiere de forma significativa. En el resto de comarcas, el p-valor obtenido es superior al nivel de significancia, por lo que no se puede rechazar la hipótesis nula y se admite que las posibles diferencias en la superficie de las naves de ambas muestras son consecuencia de la aleatoriedad.

La **Figura 3.15** comprende los diagramas de caja y bigotes de la superficie de las naves en venta y alquiler localizadas en las comarcas que se han sometido al contraste de hipótesis. En estos gráficos se puede observar la dispersión de la variable superficie según el tamaño de las cajas. En este sentido, comarcas como l'Alacantí o el Baix Vinalopó (Alicante), la Plana Alta y la Plana Baixa (Castellón) en el mercado de venta y l'Horta Nord, la Vall d'Albaida y Valencia en la provincia homónima presentan la menor dispersión. En cambio, la mayor dispersión de la superficie se da en las comarcas señaladas en el párrafo anterior, en las cuales se advierte de forma clara la diferente distribución de la superficie de las naves según el mercado al que pertenecen. Cabe destacar que, en el caso de la Hoya de Buñol en la **Figura 3.15 (c)**, la diferencia entre la superficie de alquiler y la superficie en venta es todavía mayor, tal y como recoge la **Tabla 3.11**, pero no se alcanza a ver en el diagrama por el espacio limitado para la representación.



**Figura 3.15:** Dispersión de la superficie de las naves industriales en venta y en alquiler por comarca

Por otra parte, la distribución asimétrica positiva de la superficie de las naves detectada en la [Figura 3.13](#) también se puede apreciar en estos diagramas. Así, se puede comprobar la posición de la mediana, dada por la línea gruesa, dentro de la correspondiente caja. De esta forma, aunque en líneas generales el sesgo es muy marcado y la cola positiva grande, existen conjuntos de comarca y mercado en los que puede afirmarse que la distribución de la superficie tiende a la simetría. Este hecho ocurre con las naves a la venta en el Baix Segura y en alquiler en l'Alacantí ([Figura 3.15 \(a\)](#)), así como con las naves a la venta en la Hoya de Buñol, l'Horta Oest y l'Horta Sud ([Figura 3.15 \(c\)](#)).

### 3.4. Determinación de la rentabilidad bruta

La última etapa que comprende el presente Trabajo Final de Grado consiste en hallar los índices anuales de rentabilidad bruta del alquiler de las naves industriales en el mercado de la Comunidad Valenciana, a la postre el objetivo principal que se ha definido para la realización de este estudio.

Tras haber llevado a cabo un análisis del precio y la superficie de la oferta de naves industriales de la Comunidad en función de la ubicación y haber constatado que un análisis de la rentabilidad bruta del alquiler basado en otras variables disponibles no iba a aportar resultados consistentes ni conclusiones sólidas por la escasa representatividad de dichas variables, se calculan las ratios de rentabilidad acotando las diferentes zonas geográficas disponibles: provincias, comarcas y capitales de provincia.

Los diferentes índices de rentabilidad bruta se obtienen como el RTP anual expresado en la [ecuación \(2.2\)](#), tanto en términos medios como medianos. Para ello se utiliza, por un lado, el precio medio o mediano por metro cuadrado que se debe desembolsar para adquirir una nave industrial en propiedad. Por el otro, la renta anual por metro cuadrado que debe abonarse para disfrutar de una nave en arrendamiento constituye el potencial cobro que el inversor puede recibir, de ahí su uso como estimación para el cálculo de la rentabilidad, de acuerdo con lo explicado en el [Capítulo 2](#).

En este capítulo se presentan los índices de rentabilidad bruta estimados en términos medios y medianos para las naves industriales ubicadas en la Comunidad Valenciana. Este análisis se lleva a cabo considerando diferentes zonas geográficas. En primer lugar, se ofrece una visión general de la rentabilidad que puede esperar un inversor que decida invertir en inmuebles de esta tipología para, posteriormente, exponer ratios de áreas particulares. Además, la discusión de los índices de rentabilidad bruta también se acompaña del análisis del precio por metro cuadrado tanto en el mercado de venta como en el mercado de alquiler con un doble objetivo: por una parte, contextualizar las ratios de rentabilidad obtenidas; por otra, ofrecer información adicional sobre la situación del mercado que puede resultar de interés no solo a los inversores, sino también a aquellos agentes interesados en disponer de una nave industrial para su uso, por ejemplo.

#### 4.1. Rentabilidad bruta del alquiler por provincia

En primer lugar, se analiza la situación del mercado inmobiliario de naves industriales en base a la rentabilidad por provincias. De esta manera, se dispondrá de unos indicadores iniciales respecto a los cuales podrán establecerse posteriormente comparaciones con las ratios de zonas geográficas más acotadas. La [Tabla 4.1](#) resume los índices medios de rentabilidad del alquiler, mientras que la [Tabla 4.2](#) recoge la misma información en términos medianos. En ambos casos, aparece también el precio por metro cuadrado de la oferta de naves en venta y la renta anual por metro cuadrado de las naves en alquiler.

**Tabla 4.1:** Índices de rentabilidad media por provincia

Provincia	€/m2 venta	€/m2 anuales alquiler	Rentabilidad	Naves venta	Naves alquiler
Alicante	394,44	32,25	8,18 %	455	157
Castellón	309,29	21,93	7,09 %	241	40
Valencia	379,68	33,25	8,76 %	736	371

**Tabla 4.2:** Índices de rentabilidad mediana por provincia

Provincia	€/m2 venta	€/m2 anuales alquiler	Rentabilidad	Naves venta	Naves alquiler
Alicante	348,21	28,00	8,04 %	455	157
Castellón	304,35	20,91	6,87 %	241	40
Valencia	350,00	30,00	8,57 %	736	371

De entrada, se observa que los valores medios y medianos son muy similares. Así pues, pese a que la distribución del precio y de la superficie es asimétrica, como se ha explicado en el [apartado 3.3](#), el precio por metro cuadrado se distribuye de forma más simétrica, en tanto que la media y la mediana de esta variable presentan prácticamente la misma magnitud.



La provincia de Alicante tiene el precio medio por metro cuadrado entre las naves en venta más alto, con alrededor de 395 euros por metro cuadrado, un 27 % superior al de las naves en la provincia de Castellón. En cuanto a las naves en venta, la renta media anual por unidad de superficie se sitúa en torno a los 33 euros por metro cuadrado en las provincias de Alicante y Valencia, un 50 % más que en Castellón.

Las ratios de rentabilidad también son parecidas en media y mediana. La rentabilidad media varía del 7,09 % en Castellón al 8,76 % en Valencia, mientras que la mediana oscila entre el 6,87 % y el 8,57 %. No obstante en el cálculo de la rentabilidad en la provincia de Castellón las naves en alquiler tan solo suponen una sexta parte de las naves en venta, mientras que en Valencia y Alicante representan la mitad y una tercera parte, respectivamente. Debido a esta infrarepresentación, los indicadores de rentabilidad en Castellón deberían tomarse con cautela.

## 4.2. Rentabilidad bruta del alquiler por comarca

A continuación, se realiza un análisis de la rentabilidad estimada según las distintas comarcas de la Comunidad Valenciana. Seguidamente se presentan los valores medios (Tabla 4.3) y medianos (Tabla 4.4) para su discusión.

Tabla 4.3: Índices de rentabilidad media por comarca

Comarca	€/m2 venta	€/m2 anuales alquiler	Rentabilidad	Naves venta	Naves alquiler
<b>Alicante</b>					
L'Alt Vinalopó	251,79	22,10	8,78 %	29	5
El Baix Segura	417,18	28,26	6,77 %	66	17
El Baix Vinalopó	482,88	31,22	6,47 %	80	55
El Comtat	242,38	14,23	5,87 %	26	4
L'Alacantí	527,09	38,43	7,29 %	93	35
L'Alcoià	281,13	22,30	7,93 %	33	11
La Marina Alta	403,01	46,84	11,62 %	26	10
La Marina Baixa	473,23	47,88	10,12 %	25	6
El Vinalopó Mitjà	248,02	25,14	10,14 %	77	14
<b>Castellón</b>					
El Alto Palancia	287,20	16,36	5,70 %	18	1
El Baix Maestrat	216,64	21,19	9,78 %	31	4
La Plana Alta	363,51	24,56	6,76 %	72	20
La Plana Baixa	304,01	18,98	6,24 %	120	15
<b>Valencia</b>					
El Camp de Morvedre	410,03	26,94	6,57 %	42	10
El Camp del Túria	400,92	40,67	10,14 %	89	60
La Hoya de Buñol	281,11	46,92	16,69 %	16	18
L'Horta Nord	439,52	34,32	7,81 %	59	22
L'Horta Oest	401,97	34,16	8,50 %	181	109
L'Horta Sud	420,90	34,43	8,18 %	86	73
La Canal de Navarrés	190,53	12,00	6,30 %	6	1
La Costera	271,48	20,82	7,67 %	30	7
La Plana de Utiel-Requena	203,28	16,20	7,97 %	6	2
La Vall d'Albaida	223,01	15,09	6,77 %	48	9
La Ribera Alta	327,59	20,63	6,30 %	66	20
La Ribera Baixa	369,34	22,69	6,14 %	25	13
La Safor	353,27	28,67	8,12 %	60	15
<b>Valencia</b>	<b>631,68</b>	<b>27,22</b>	<b>4,31 %</b>	<b>22</b>	<b>12</b>

Al segmentar por comarcas se aprecia una mayor desviación de la media con respecto a la mediana en el precio por metro cuadrado tanto de la oferta de naves en venta como la del alquiler. En consecuencia, los índices medios y medianos de rentabilidad en algunas comarcas difieren de forma significativa, por ejemplo en l'Alt Vinalopó (Alicante) o la Hoya de Buñol y la Plana de Utiel-Requena (Valencia). Sin embargo, también es cierto que en estas zonas la oferta de naves es industriales es escasa o sus características son muy particulares, como se discute en el [apartado 3.3](#).

En Alicante, los precios por metro cuadrado se dan en las comarcas del cono sur en las que, por otra parte se concentra la oferta. L'Alacantí es la comarca con el precio por unidad de superficie de nave industrial en venta más caro, mientras que las comarcas de la Marina se sitúan a la cabeza en la renta anual por metro cuadrado. Por otro lado, la Plana Alta es la comarca castellanense con el precio más alto.

Respecto a la provincia de Valencia, la comarca homónima es la que presenta el precio por unidad de superficie de venta más elevado, superior a los 600 euros por metro cuadrado. El Camp de Morvedre, el Camp del Túria y el área metropolitana de Valencia (las comarcas de l'Horta) son las zonas que también presentan un precio medio y mediano de venta superior al de la provincia.

**Tabla 4.4:** Índices de rentabilidad mediana por comarca

Comarca	€/m2 venta	€/m2 anuales alquiler	Rentabilidad	Naves venta	Naves alquiler
<b>Alicante</b>					
L'Alt Vinalopó	226,67	25,34	11,18 %	29	5
El Baix Segura	407,75	25,20	6,18 %	66	17
El Baix Vinalopó	430,95	28,00	6,50 %	80	55
El Comtat	214,00	9,88	4,62 %	26	4
L'Alacantí	493,63	35,29	7,15 %	93	35
L'Alcoià	215,05	19,29	8,97 %	33	11
La Marina Alta	397,72	44,75	11,25 %	26	10
La Marina Baixa	500,00	48,44	9,69 %	25	6
El Vinalopó Mitjà	233,10	21,13	9,06 %	77	14
<b>Castellón</b>					
El Alto Palancia	292,06	16,36	5,60 %	18	1
El Baix Maestrat	194,44	16,49	8,48 %	31	4
La Plana Alta	346,30	24,13	6,97 %	72	20
La Plana Baixa	295,58	17,91	6,06 %	120	15
<b>Valencia</b>					
El Camp de Morvedre	364,21	25,74	7,07 %	42	10
El Camp del Túria	359,34	42,59	11,85 %	89	60
La Hoya de Buñol	243,26	51,60	21,21 %	16	18
L'Horta Nord	444,44	32,22	7,25 %	59	22
L'Horta Oest	374,61	29,27	7,81 %	181	109
L'Horta Sud	403,07	33,00	8,19 %	86	73
La Canal de Navarrés	159,96	12,00	7,50 %	6	1
La Costera	236,25	21,49	9,10 %	30	7
La Plana de Utiel-Requena	150,85	16,20	10,74 %	6	2
La Vall d'Albaida	188,33	16,80	8,92 %	48	9
La Ribera Alta	273,12	20,00	7,32 %	66	20
La Ribera Baixa	360,39	18,05	5,01 %	25	13
La Safor	328,33	28,00	8,53 %	60	15
Valencia	613,68	23,71	3,86 %	22	12

En el extremo opuesto se sitúan la Canal de Navarrés, la Plana de Utiel y la Vall d'Albaida. La situación se replica en el mercado de alquiler con la salvedad de la comarca de la capital, en la que la renta anual por unidad de superficie es incluso inferior a la media de la provincia.

En cuanto a los índices de rentabilidad, las comarcas donde la ratio es más alta son aquellas en las que se cuestionó en el apartado 3.3 su validez por la disparidad entre las superficies disponibles a la venta y en alquiler. Son, en términos medios, las comarcas de la Marina en Alicante y el Camp del Túria y la Hoya de Buñol en Valencia. En Alicante, las rentabilidades estimadas se sitúan por debajo de la media provincial a excepción del Vinalopó Mitjà de modo que, obviando las observaciones de la Marina, estos indicadores parecen más fiables. En Castellón, la rentabilidad más alta se da en el Baix Maestrat, pero el escaso número de naves utilizadas en el cálculo resta fiabilidad a dicho valor. Así pues, el casi 7 % tanto de media como en términos medianos de la Plana Alta es el valor más alto, aunque la Plana Baixa tan solo anota medio punto y un punto porcentual menos, respectivamente.

En la provincia de Valencia, si se obvian las comarcas con una rentabilidad artificialmente alta por las particularidades del mercado de alquiler, esto es, naves de gran superficie, resulta que la rentabilidad de todas las comarcas restantes se mueve por debajo de la media y la mediana de la provincia. En este sentido, al igual que en Alicante, estas ratios seguramente son más representativas de la rentabilidad que puede esperar un inversor. Sí cabe destacar la rentabilidad estimada para la comarca de Valencia, la menor de toda la Comunidad. El alto precio por metro cuadrado de las naves en venta y una renta anual por metro cuadrado en el extremo inferior del rango penalizan el retorno de la inversión.

Como resumen, se incluye en la Figura 4.1 un mapa con las rentabilidades brutas medias recogidas en la Tabla 4.3.

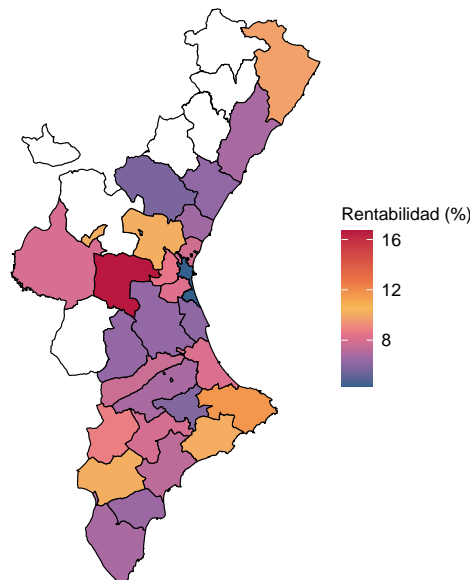


Figura 4.1: Mapa de rentabilidad media del alquiler de las naves industriales por comarca

A lo largo del estudio se ha constatado que la oferta de naves industriales en las tres provincias se configura de forma distinta en las zonas costeras y en las de interior. Por ello, también se analiza la rentabilidad en cada provincia agrupando sus correspondientes comarcas en función de su localización con respecto al mar. La [Tabla 4.5](#) resume los indicadores medios de precio y rentabilidad bruta de alquiler, mientras que la [Tabla 4.6](#) recoge los valores medianos de estas variables.

**Tabla 4.5:** Índices de rentabilidad media por comarcas de costa e interior

Provincia	Zona	€/m <sup>2</sup> venta	€/m <sup>2</sup> anuales alquiler	Rentabilidad	Naves venta	Naves alquiler
Alicante	Costa	474,11	34,94	7,37 %	290	123
	Interior	254,42	22,49	8,84 %	165	34
Castellón	Costa	311,08	22,07	7,09 %	223	39
	Interior	287,20	16,36	5,70 %	18	1
Valencia	Costa	402,76	33,43	8,30 %	630	334
	Interior	242,54	31,64	13,05 %	106	37

**Tabla 4.6:** Índices de rentabilidad mediana por comarcas de costa e interior

Provincia	Zona	€/m <sup>2</sup> venta	€/m <sup>2</sup> anuales alquiler	Rentabilidad	Naves venta	Naves alquiler
Alicante	Costa	445,22	31,51	7,08 %	290	123
	Interior	226,42	20,87	9,22 %	165	34
Castellón	Costa	306,51	21,00	6,85 %	223	39
	Interior	292,06	16,36	5,60 %	18	1
Valencia	Costa	370,03	30,21	8,16 %	630	334
	Interior	214,25	27,10	12,65 %	106	37

La costa alicantina presenta el precio por metro cuadrado de las naves en venta más alto, un 86 % superior de media y casi el doble en términos medios que los inmuebles situados en el interior de la provincia. En cambio, el precio de venta más bajo corresponde a las naves del interior de Valencia, alrededor de un 70 % inferiores a los inmuebles industriales situados en comarcas costeras. En Castellón, en cambio, no existe una diferencia notable en el precio por metro cuadrado del mercado de venta. En el segmento de alquiler, por su parte, la renta anual por metro cuadrado más alta se alcanza en la costa de Valencia y Alicante, con valores muy similares. Asimismo, en este caso es en Valencia donde apenas varía el precio entre las comarcas costeras y las de interior.

Si se presta atención a las ratios de rentabilidad bruta, vuelve a destacar el interior de Valencia. Ello se debe a que la comarca de la Hoya de Buñol, en la que se da la particular situación de naves grandes y precios altos en el segmento de alquiler y naves de superficie más contenida y precios más moderados en el mercado de venta, concentra la mayor parte de la oferta del interior de la provincia. Sin embargo, en la provincia de Alicante, cuyos mercados de venta y alquiler son más comparables, también se da la misma situación, siendo la rentabilidad en el interior superior a la media y mediana, e inferior en la costa. En cuanto a Castellón, la rentabilidad en la costa es prácticamente igual a la del conjunto de la provincia puesto que tan solo consta una nave en alquiler en comarcas del interior, de forma que el dato de rentabilidad correspondiente carece de utilidad.

### 4.3. Rentabilidad bruta del alquiler en las capitales de provincia

La última parte del estudio de la rentabilidad del alquiler de naves industriales comprende el mercado inmobiliario de las capitales de provincia, a excepción de Alicante, junto a la que también se considera la ciudad de Elche. En la [Tabla 4.7](#) se presentan los resultados medios, mientras que la [Tabla 4.8](#) contiene los valores medianos. Estos indicadores pueden ser comparados con los obtenidos para las comarcas de referencia a fin de obtener conclusiones.

**Tabla 4.7:** Índices de rentabilidad media por capitales de provincia

Ciudad	€/m2 venta	€/m2 anuales alquiler	Rentabilidad	Naves venta	Naves alquiler
Alicante / Elche	518,65	35,43	6,83 %	104	75
Castellón de la Plana	406,39	28,75	7,07 %	23	9
Valencia	617,21	27,65	4,48 %	19	10

**Tabla 4.8:** Índices de rentabilidad mediana por capitales de provincia

Ciudad	€/m2 venta	€/m2 anuales alquiler	Rentabilidad	Naves venta	Naves alquiler
Alicante / Elche	451,03	32,86	7,29 %	104	75
Castellón de la Plana	373,36	26,44	7,08 %	23	9
Valencia	656,34	22,07	3,36 %	19	10

Si se atiende al precio por metro cuadrado, el correspondiente a las naves en venta es sensiblemente inferior en Alicante con respecto a la media de la comarca, mientras que se puede considerar igual en Valencia. Ello confirma lo refrendado en la [Tabla 3.5](#) sobre el predominio de las naves situadas en las capitales de estas provincias con respecto a sus comarcas de cabecera. No obstante, en Castellón de la Plana el precio medio es un 11 % inferior a la media de la comarca de la Plana Alta, no así el precio mediano, que experimenta un crecimiento del 8 %. Del mismo modo, estos valores confirman lo descubierto al analizar la distribución geográfica de la oferta de inmuebles ya que, a diferencia de las ciudades anteriores, las naves a la venta en Castellón de la Plana no son mayoría dentro de la comarca. Asimismo, estos apuntes sobre el mercado de venta aplican también al segmento de alquiler.

En términos de rentabilidad, es evidente que la tendencia del precio por metro cuadrado de las naves se reproduce en las ratios correspondientes. Así pues, la rentabilidad es, de media, 0,5 puntos inferior en Alicante y Elche en comparación con la comarca de l'Alacantí, mientras que se mantiene si se observa la mediana. En Castellón de la Plana, sube tres décimas de media y una décima en términos medianos con respecto a la comarca de la Plana Alta. Una evolución similar se registra también en Valencia capital con respecto a la comarca homónima.

A lo largo de este Trabajo Final de Grado se ha estudiado el mercado inmobiliario de naves comerciales en la Comunidad Valenciana para determinar la rentabilidad que un inversor puede esperar de una inversión en este tipo de inmueble. Se considera que disponer de esta clase de información aporta mucho valor en la toma de decisiones de inversión, en tanto que permite ampliar el abanico de alternativas de inversión a un segmento del mercado inmobiliario que, a veces, no recibe la misma atención que otros inmuebles.

El trabajo ha consistido, en primer lugar, en contextualizar el mercado inmobiliario de naves industriales en España y, más concretamente, en la Comunidad Valenciana, en la que se centra el estudio. Se han comentado las diferencias que manifiestan las naves industriales y su mercado con respecto a las viviendas, la tipología de inmueble predominante, y los locales comerciales, por su condición de inmuebles destinados al uso empresarial.

En segundo lugar, se han repasado los conceptos de inversión y rentabilidad, básicos para la comprensión y el desarrollo del trabajo. En cuanto a la primera, se han descrito los tipos de inversión que existen y se ha acotado la definición de inversión inmobiliaria, dejando claro qué inversiones pueden acogerse a dicha definición y cuáles no. Por otro lado, se ha explicado qué se entiende por rentabilidad, los distintos tipos de rentabilidad que pueden calcularse y las ventajas e inconvenientes de los mismos. En relación a ello, se ha motivado el uso del índice de rentabilidad bruta del alquiler o rent-to-price (RTP) como ratio de referencia para cumplir el objetivo principal de este trabajo.

A continuación, ha sido necesario un proceso de aprendizaje para manejar y aprovechar todas las posibilidades que ofrece el lenguaje de programación R. Este lenguaje de programación constituye la base del estudio llevado a cabo en el presente trabajo. La obtención de una muestra de datos, su depuración, el análisis estadístico y la propia confección del documento (gráficas, tablas, mapas) se han realizado utilizando R y las distintas librerías de código abierto disponibles.

Para disponer de una muestra de la oferta de naves industriales a la venta y en alquiler se ha implementado en R un *web scraper*, una herramienta automatizada capaz de extraer información de sitios web. Para ello, previamente se ha realizado un ejercicio de comprensión sobre la técnica de *web scraping* y cómo se pueden aprovechar los datos disponibles en Internet a nuestro favor. Este era uno de los objetivos marcados al iniciar el trabajo y puede afirmarse que se ha cumplido con creces, pues se dispone de un *software* perfectamente preparado para obtener información de los anuncios de naves industriales disponibles en un popular portal inmobiliario.

Los datos extraídos se han manipulado para disponer de la máxima información y que las muestras de la oferta de naves en venta y en alquiler fueran representativas del mercado inmobiliario. Con el primer objetivo en mente, se han reestructurado los datos y se han estudiado y aplicado técnicas propias de la minería de textos o *text mining* para descubrir términos repetidos en los

textos de los anuncios y convertirlos en variables descriptivas de los inmuebles. Si bien se han empleado con éxito estas técnicas, las variables de nueva generación no se han utilizado en fases posteriores del trabajo debido a que gran parte de las mismas son variables cualitativas binarias y no se ha podido decidir con rigurosidad si no haberlas registrado implica que el inmueble no presenta un atributo determinado.

La base de datos definitiva para el estudio de la rentabilidad se ha logrado tras depurar las observaciones originales de las muestras suprimiendo aquellos valores anómalos u *outliers* que diferían en exceso del conjunto central. Para ello, se han aplicado tanto técnicas univariantes como multivariantes cuidando terminar con una muestra homogénea del mercado de naves comerciales, pero que al mismo tiempo no perdiera representatividad y ciertas clases de naves que también forman parte de dicho mercado.

Con una base de datos consistente se ha llevado a cabo un análisis estadístico de la oferta de naves en venta y en alquiler en la Comunidad Valenciana centrado en tres variables: la localización, la superficie y el precio. Las conclusiones más interesantes sobre las dos primeras son las siguientes:

- Las muestras son desiguales, puesto que alrededor del 70 % de las naves están a la venta y solo el 30 % se ofertan en alquiler.
- La mayor parte de las naves, independientemente del mercado, se sitúan en la costa: el sur de la provincia de Alicante, el área metropolitana de Valencia y el norte de dicha provincia y las comarcas de la Plana Alta y Baixa en Castellón.
- El peso de las naves ubicadas en las capitales de provincia es significativo con respecto a las comarcas de referencia en Alicante y Valencia, no así en Castellón.
- La superficie media de las naves en venta es de casi 1.000 metros cuadrados, mientras que las ofertadas en alquiler alcanzan, de media, los 1.940 metros cuadrados.
- Sin embargo, el 50 % de las naves a la venta tiene menos de 723 metros cuadrados; en el segmento de alquiler, la cifra asciende a 900 metros cuadrados.
- A grandes rasgos, no existen diferencias significativas en la distribución de la superficie de las naves por comarcas, si bien destacan con valores más altos la Marina (Alicante), con una superficie media de las naves en venta de 1.500 metros cuadrados, y la Hoya de Buñol, el Camp de Morvedre y el Camp del Túria, donde se ofertan en alquiler grandes naves logísticas con una superficie media de 12.000 metros cuadrados en la primera comarca y 2.600 metros cuadrados en las dos últimas.
- Se han detectado comarcas en las que la distribución de la superficie entre los mercados de venta y alquiler no se asemeja y que invalidan las conclusiones que se puedan alcanzar sobre la rentabilidad, al tratarse en definitiva de muestras de naves diferentes: la Marina Baixa y Alta, el Camp de Morvedre, el Camp del Túria, la Hoya de Buñol, l'Horta Oest, l'Horta Sud y la Ribera Baixa.

Finalmente, se han determinado los precios por metro cuadrado de las naves en venta y las rentas anuales por metro cuadrado de las ofertadas en alquiler y se han obtenido los índices de rentabilidad bruta anual. Los aspectos más destacados son los siguientes:

- Por provincias, Alicante y Valencia comparten el mismo precio mediano por metro cuadrado en los mercados de venta y alquiler (350 euros por metro cuadrado y 30 euros anuales por metro cuadrado), resultando en una rentabilidad media y mediana por encima del 8 %. En Castellón, la menor renta anual por metro cuadrado (20 euros), reduce la rentabilidad al entorno del 7 %.
- Por comarcas, algunas de los precios por metro cuadrado y rentabilidades más altas se dan en las comarcas señaladas anteriormente para ser descartadas.
- Si se considera una división territorial por comarcas, L'Alacantí presenta el precio medio por metro cuadrado de venta más alto (527 euros), mientras que el Baix Maestrat y la Vall d'Albaida son las zonas con el precio más bajo, en el entorno de los 200 euros por metro cuadrado.
- Las rentabilidades varían ampliamente en términos medios, desde el 4,31 % en la comarca de Valencia capital al 10,14 % en el Vinalopó Mitjà alicantino.
- En líneas generales y dada la distribución marcadamente asimétrica del precio y la superficie, la mediana es un estadístico más robusto y la rentabilidad mediana una ratio que refleja con mayor precisión el retorno esperado.
- A pesar de que los precios más altos se dan en las naves situadas en comarcas costeras, la rentabilidad estimada es superior en las comarcas de interior, debido a la importante corrección en el precio de las naves en venta.
- Si se analiza la rentabilidad de las naves en las capitales de provincia, se mantiene con respecto a la comarca de referencia en el entorno del 7 % o disminuye ligeramente.

Los resultados que se han obtenido en este trabajo pueden variar ligeramente en función de las muestras que se considere. Por un lado, no hay motivos que conduzcan a pensar que las muestras utilizadas no son representativas del mercado de naves industriales de la Comunidad Valenciana. Sin embargo, sí es cierto que algunas zonas de interior están notablemente infrarrepresentadas y debería evaluarse si la cantidad de anuncios de un portal inmobiliario muestra con precisión el tamaño del mercado en esas zonas.

En la misma línea, también podría considerarse una revisión de los métodos empleados para eliminar observaciones anómalas en aras de tener una muestra más homogénea y menos representativa, si resultase de interés para la consecución de los objetivos del estudio. Finalmente, si la estructuración de los datos en origen fuese más consistente, se podría haber llevado a cabo un estudio más completo considerando otras variables descriptivas de interés.

Como trabajos futuros que podrían desarrollarse a partir de este, se pueden plantear varios temas dado que se dispone de una potente herramienta de *web scraping* y una metodología de análisis de la información. Puesto que este estudio de la rentabilidad es una imagen fija, sería interesante monitorizar el mercado de forma periódica y contar con indicadores de precio y rentabilidad lo más actualizados posibles. Asimismo, se podría estudiar un mercado inmobiliario distinto, por ejemplo, el de la vivienda o los locales comerciales. También se podrían utilizar los datos del mercado con otra finalidad, como la de diseñar un modelo predictivo del precio.





# Bibliografía

---

- Amargant, R. (2008). **La inversión en productos inmobiliarios: el mercado inmobiliario, la inversión en inmuebles, vehículos para la inversión y planificación inmobiliaria** (Vol. 5). Profit Editorial. (Vid. pág. 1).
- Aranda, J. L. (2021). **La edad dorada de las grandes naves industriales**. *El País*. <https://elpais.com/economia/2021-09-27/la-edad-dorada-de-las-grandes-naves-industriales.html> (vid. pág. 5).
- Banco Central Europeo. (2022). El tipo de interés negativo del BCE. <https://www.ecb.europa.eu/ecb/educational/explainers/tell-me-more/html/why-negative-interest-rate.es.html> (vid. pág. 1).
- Banco de España. (2022). **Indicadores del mercado inmobiliario**. [https://www.bde.es/webbde/es/estadis/infoest/si\\_1\\_5.pdf](https://www.bde.es/webbde/es/estadis/infoest/si_1_5.pdf) (vid. pág. 5).
- Brotos Martínez, J. M. (2017). **Supuestos de valoración de inversiones**. Universidad Miguel Hernández. (Vid. págs. 10-12).
- Butler, C. (1985). **Statistics in linguistics**. Blackwell Publishers. (Vid. pág. 44).
- Casas del Rosal, J. C., Caridad y Ocerín, J. M., & Núñez Tabales, J. M. (2018). **La aglomeración urbana de la ciudad de Sevilla: análisis de la oferta de inmuebles destinados a la actividad comercial e industrial**. *Revista de métodos cuantitativos para la economía y la empresa*, 25, 72-92 (vid. pág. 1).
- Casas-Rosal, J. C., Núñez-Tabales, J. M., & Caridad, L. (2018). **Una mirada al mercado de locales comerciales a la venta de la ciudad de Córdoba**. *International Journal of Scientific Management and Tourism*, 4, 61-71 (vid. págs. 6, 18).
- Cervellera, A. (2021). **La compraventa de inmuebles repunta un 11,8 % por el interés en naves y locales comerciales**. *Las Provincias*. <https://www.lasprovincias.es/economia/compraventa-inmuebles-aumenta-20210111190958-nt.html> (vid. pág. 6).
- Contreras, I. (2006). **Análisis de la rentabilidad económica (ROI) y financiera (ROE) en empresas comerciales y en un contexto inflacionario**. *Visión Gerencial*, 13-28 (vid. pág. 11).
- Diouf, R., Sarr, E. N., Sall, O., Birregah, B., Bousso, M., & Mbaye, S. N. (2019). **Web scraping: state-of-the-art and areas of application**. *2019 IEEE International Conference on Big Data (Big Data)*, 6040-6042 (vid. págs. 13, 14).

- Fabra Garcés, L. A., Salvador Figueras, M., & Gargallo Valero, P. (2021). **Estadística registral inmobiliaria. Anuario 2021**. Colegio de Registradores de la Propiedad, Bienes Muebles y Mercantiles de España. (Vid. págs. 4-6).
- Fernández Cerezo, A., Lamas, M., Roibás, I., & Vegas, R. (2021). **El impacto de la pandemia de COVID-19 en el mercado inmobiliario comercial español**. *Boletín Económico/Banco de España*, 3 (vid. págs. 2, 3).
- Glez-Peña, D., Lourenço, A., López-Fernández, H., Reboiro-Jato, M., & Fdez-Riverola, F. (2014). **Web scraping technologies in an API world**. *Briefings in Bioinformatics*, 15, 788-797 (vid. págs. 14, 15).
- Hassler, M., & Fliedl, G. (2006). **Text preparation through extended tokenization**. *WIT Transactions on Information and Communication Technologies*, 37, 13-21 (vid. pág. 24).
- Hotho, A., Nürnberger, A., & Paaß, G. (2005). **A brief survey of text mining**. *LDV Forum - GLDV Journal for Computational Linguistics and Language Technology* (vid. pág. 24).
- Idealista. (2020). La rentabilidad de la inversión en vivienda se estanca en 2019 en el 7,4 %. <https://www.idealista.com/news/inmobiliario/vivienda/2020/01/21/779550-la-rentabilidad-de-la-inversion-en-vivienda-se-estanca-en-2019-en-el-7-4> (vid. pág. 2).
- Idealista. (2022a). idealista — Casas y pisos, alquiler y venta. Anuncios gratis. <https://www.idealista.com/> (vid. págs. 19, 20).
- Idealista. (2022b). La rentabilidad de la inversión en vivienda cae hasta el 6,9 % en 2021. <https://www.idealista.com/news/inmobiliario/vivienda/2022/01/10/794200-la-rentabilidad-de-la-inversion-en-vivienda-cae-hasta-el-6-9-en-2021> (vid. pág. 5).
- Instituto de Comercio Exterior. (s.f.). Comunidad Valenciana - Industrias destacadas. <https://www.investinspain.org/content/icex-invest/es/regions/comunidad-valenciana/industrias-destacadas.html> (vid. pág. 6).
- Junta de Normas Internacionales de Contabilidad. (2018). *NIC 40 Propiedades de Inversión*. (Vid. pág. 10).
- Krotov, V., Johnson, L. R., & Silva, L. (2020). **Tutorial: legality and ethics of web scraping**. *Communications of the Association for Information Systems*, 47, 539-563 (vid. págs. 13-16).
- Lamas, M., & Romaniega, S. (2022). **Elaboración de un índice de precios para el mercado inmobiliario comercial de España**. *Documentos Ocasionales/Banco de España*, 2203 (vid. págs. 2, 3).
- Leys, C., Delacre, M., Mora, Y. L., Lakens, D., & Ley, C. (2019). **How to classify, detect, and manage univariate and multivariate outliers, with emphasis on pre-registration**. *International Review of Social Psychology*, 32 (vid. pág. 27).
- Maesschalck, R. D., Jouan-Rimbaud, D., & Massart, D. L. (2000). **The Mahalanobis distance**. *Chemometrics and Intelligent Laboratory Systems*, 50, 1-18 (vid. pág. 28).
- Oi Realtor. (2021). La demanda de naves industriales en España. <https://www.oirealtor.com/noticias-inmobiliarias/la-demanda-de-naves-industriales-en-espana/> (vid. pág. 6).

- Porter, M. F. (1980). **An algorithm for suffix stripping**. *Program: electronic library and information systems*, 14, 130-137 (vid. pág. 25).
- Raisin. (s.f.). Rentabilidad: qué es, tipos y cómo se calcula. <https://www.raisin.es/educacion-financiera/rentabilidad-que-es-tipos-y-como-se-calcula/> (vid. pág. 11).
- Rousseeuw, P. J., & Hubert, M. (2011). **Robust statistics for outlier detection**. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, 1, 73-79 (vid. pág. 27).
- Rousseeuw, P. J., & van Zomeren, B. C. (1990). **Unmasking multivariate outliers and leverage points**. *Journal of the American Statistical Association*, 85, 633-639 (vid. págs. 28, 29).
- SimilarWeb. (2022). Clasificación de los principales sitios web para inmobiliaria en España en julio 2022. <https://www.similarweb.com/es/top-websites/spain/category/business-and-consumer-services/real-estate/> (vid. pág. 18).
- Singh, P. (2016). **Investment Management**. Himalaya Publishing House. (Vid. págs. 9, 10).
- Yao, Z., & Ze-wen, C. (2011). **Research on the construction and filter method of stop-word list in text preprocessing**. *2011 Fourth International Conference on Intelligent Computation Technology and Automation (ICICTA)*, 1, 217-221 (vid. pág. 24).
- Zhao, B. (2022). **Web Scraping** (L. A. Schintler & C. L. McNeely, Eds.). Springer International Publishing. (Vid. págs. 13, 14, 16).



# A

## Objetivos de Desarrollo Sostenible

**Grado de relación del TFG con los Objetivos de Desarrollo Sostenible de la Agenda 2030**



Objetivos de Desarrollo Sostenible	Alto	Medio	Bajo	No procede
ODS 1. Fin de la pobreza				X
ODS 2. Hambre cero				X
ODS 3. Salud y bienestar				X
ODS 4. Educación de calidad				X
ODS 5. Igualdad de género				X
ODS 6. Agua limpia y saneamiento				X
ODS 7. Energía asequible y no contaminante				X
ODS 8. Trabajo decente y crecimiento económico		X		
ODS 9. Industria, innovación e infraestructuras	X			
ODS 10. Reducción de las desigualdades				X
ODS 11. Ciudades y comunidades sostenibles		X		
ODS 12. Producción y consumo responsables				X
ODS 13. Acción por el clima				X
ODS 14. Vida submarina				X
ODS 15. Vida de ecosistemas terrestres				X
ODS 16. Paz, justicia e instituciones sólidas				X
ODS 17. Alianzas para lograr objetivos				X

En este anexo se analiza la relación entre el Trabajo Final de Grado desarrollado con los diferentes Objetivos de Desarrollo Sostenible (ODS); establecidos por la Organización de las Naciones Unidas (ONU).

En el año 2015, los países miembros de las Naciones Unidas firmaron una serie de 17 objetivos con el fin de erradicar la pobreza, proteger el planeta y alcanzar la prosperidad entre todos. Para cumplir con los objetivos, se han definido metas con el propósito de satisfacerlas antes de 2030, con la ayuda tanto de los gobiernos como del sector público y privado, y lo más importante, con la ayuda de todos nosotros.

Analizando los diferentes objetivos existentes, partimos del **ODS 9** sobre la **industria, innovación e infraestructuras**. La ciencia de datos y las técnicas que se emplean en este TFG para extraer datos y obtener información es un campo de innovación con mucho recorrido, que simplifica el trabajo y potencia la productividad frente a técnicas más tradicionales. Ello se hace todavía más patente si se trabaja con gran cantidad de datos y se tienen en cuenta muchas variables en el proceso de análisis.

Además, otro de los objetivos que se puede relacionar con este trabajo es el **ODS 11**, que hace referencia a las **ciudades y comunidades sostenibles**. Teniendo en cuenta la meta 11.1, dirigida a asegurar el acceso de todas las personas a viviendas seguras y asequibles, teniendo en cuenta la inflación existente y la tendencia sostenida al alza de los alquileres en todos los sectores inmobiliarios, una herramienta de *web scraping* como la desarrollada en este trabajo resulta muy útil para analizar las diferentes plataformas existentes, facilitando así el estudio del mercado inmobiliario y la accesibilidad al mismo, dada la escasa oferta desde la pandemia provocada por el COVID-19 y la alta demanda existente.

Para finalizar, cabe hacer énfasis en la importancia que tiene el análisis de datos en el mercado laboral en la actualidad, y cómo apostar por la Inteligencia de Negocios en la docencia universitaria y en las empresas otorga a muchos jóvenes la oportunidad de aprender las bases necesarias de una herramienta que será y ya está siendo demandada en diversos sectores del mercado laboral. Con ello, la Inteligencia de Negocios y sus aplicaciones pueden relacionarse con el **ODS 8**, que engloba el **trabajo decente y crecimiento económico**.