

# Grado en Estadística

---

**Título: Precios de viviendas, regulación y desigualdades económicas**

**Autor: Albert Canadell**

**Director: M. Guillen, A. Estruch**

**Departamento: Econometría, Estadística y Economía aplicada**

**Convocatoria: Junio 2016**

## Resumen:

El mercado de la vivienda español es un mercado ineficiente, y como consecuencia muchas personas no tienen acceso a la vivienda. En este trabajo se expone la situación actual del mercado de la vivienda junto con breves referencias a la situación en décadas pasadas para luego, tratar las variables del Atlas Digital, una base de datos del Ministerio de Fomento con datos relacionados con la vivienda. Estos datos son tratados a través de medidas de desigualdad (estadísticos e índices) y también a través de la realización de modelos de regresión para entender el comportamiento de algunas de las variables que se ofrecen en el Atlas Digital para determinar posibles soluciones a las ineficiencias del mercado de la vivienda.

## Abstract:

*The Spanish housing market is inefficient. As a result, many people have no access to housing. This paper aims to explain the current situation of the housing market with brief references of the situation in past decades.*

*This paper treats variables related to housing from a public access database, property of the Ministry of Public Works and Transport. After a first exploratory analysis, data is processed through inequality measures (statistics and indices), and through regression models to understand the behaviour of some of the variables that are offered in the database to determine possible solutions to the inefficiencies found in the housing market.*

*Keywords:* “Mercado de la vivienda”, “Parque de vivienda”, “Ocupación del suelo”, ”Desigualdad”, “Gini”, ”Área Urbana”, ”Propiedad”, “Alquiler”.

## Clasificación AMS

- 62J05 Linear Regression.
- 65D18 Computer graphics, image analysis, and computational geometry.

## ÍNDICE

<b>I.</b>	<b>INTRODUCCIÓN Y METODOLOGIA.....</b>	<b>1</b>
<b>II.</b>	<b>SITUACIÓN ACTUAL DEL MERCADO DE LA VIVIENDA.....</b>	<b>2</b>
<b>2.1.</b>	<b>Situación actual en España .....</b>	<b>2</b>
2.1.1.	Mercado de la vivienda .....	5
2.1.2.	Coste de la vivienda.....	13
<b>2.2.</b>	<b>Problemas del mercado de la vivienda español .....</b>	<b>14</b>
<b>III.</b>	<b>DATOS DE ESTUDIO .....</b>	<b>16</b>
<b>3.1.</b>	<b>Fuente de datos. Atlas Digital.....</b>	<b>16</b>
3.1.1.	Funcionamiento del Atlas Digital.....	16
<b>3.2.</b>	<b>Variables de estudio .....</b>	<b>18</b>
3.2.1.	Variables demográficas .....	19
3.2.2.	Variables del mercado de la vivienda.....	19
3.2.3.	Variables del parque de vivienda .....	20
3.2.4.	Variables de la ocupación del suelo .....	21
<b>3.3.</b>	<b>Descriptiva de los datos .....</b>	<b>21</b>
3.3.1.	Descriptiva por municipios .....	22
3.3.2.	Descriptiva por Comunidades Autónomas.....	23
<b>3.4.</b>	<b>Tratamiento de datos <i>missing</i> y interpolación y extrapolación de datos.....</b>	<b>45</b>
<b>3.5.</b>	<b>Descriptiva de los datos tratados .....</b>	<b>46</b>
<b>IV.</b>	<b>TRATAMIENTO DE LAS DESIGUALDADES A TRAVÉS DE ÍNDICES .....</b>	<b>51</b>
4.1.	¿Qué se entiende por desigualdad? Cómo medirla a través de estadísticos e índices ...	51
4.2.	Definición de estadísticos e índices. Propiedades .....	51
4.3.	Aplicación de los métodos de medición presentados a los datos de estudio.....	55
<b>5.</b>	<b>MODELO DE PANEL PARA EXPLICAR EL COMPORTAMIENTO DEL PRECIO DE LA VIVIENDA NUEVA .....</b>	<b>63</b>
<b>5.1.</b>	<b>Definición del modelo .....</b>	<b>63</b>
5.1.1.	Supuestos del modelo .....	64
5.1.2.	Estimación de los parámetros poblacionales. ....	64
5.1.3.	Bondad del ajuste.....	65
5.1.4.	Especificaciones del modelo y diagnósticos.....	65
<b>5.2.</b>	<b>Análisis de los datos de panel.....</b>	<b>72</b>
<b>5.3.</b>	<b>Validación del modelo.....</b>	<b>75</b>
5.3.1.	Test de heterogeneidad .....	75
5.3.2.	Supuesto de homoscedasticidad.....	75
5.3.3.	Efectos fijos vs efectos aleatorios .....	75
<b>6.</b>	<b>SOLUCIONES Y CONCLUSIONES .....</b>	<b>76</b>
	<b>BIBLIOGRAFÍA .....</b>	<b>79</b>

<b>ANEXO .....</b>	<b>81</b>
<b>A.1 Definición de variables.....</b>	<b>81</b>
<b>A.2. Ejemplos del código implementado en R .....</b>	<b>83</b>
A.2.1. Carga de datos y de los paquetes necesarios.....	83
A.2.2. Lectura de datos.....	84
A.2.3. Imputación de <i>missings</i> .....	84
A.2.4. Interpolación de datos .....	84
A.2.5. Creación de las bases de datos anuales .....	85
A.2.6. Creación de la base de datos de panel.....	85
A.2.7. Ejemplos de gráficos del trabajo .....	85
A.2.8. Estimación del modelo con datos de panel.....	87
A.2.9. Cálculo de la desigualdad .....	87

## I. INTRODUCCIÓN Y METODOLOGIA

El mercado de la vivienda español es un mercado basado en la vivienda en propiedad. Esta forma de acceso a la vivienda ha ayudado a la economía nacional a crecer pero, sin embargo, cuando la economía entra en recesión como es el caso de la situación actual, el acceso a la vivienda pasa a ser un problema para muchas familias. De aquí, que hayan aparecido movimientos sociales como la Plataforma de Afectados por la Hipoteca, surgida en 2009, con la finalidad de dar soporte a las familias que no pueden seguir pagando hipotecas que firmaron en época de expansión, cuando los ingresos familiares eran representativamente superiores a los gastos.

Este trabajo no entrará en el debate sobre la actuación del sector bancario español, sobre si los bancos deberían asumir responsabilidades sobre las hipotecas firmadas y por los riesgos tomados, sino que se centrará en las distintas formas de acceso a la vivienda. Para ello, este trabajo pretende estudiar la composición del mercado de la vivienda español y su evolución a lo largo de los últimos años, para detectar las ineficiencias de este y cuantificarlas. Además, intentará plantear posibles soluciones, intentando que estas puedan ser aplicadas a corto o medio plazo, dado que la situación económica actual lo requiere.

Los datos de estudio se toman del Atlas Digital, un geoportal del Ministerio de Fomento con datos referentes a la vivienda de los últimos años, que permite descargar los datos en formato Excel.

Los datos en formato Excel son tratados en un primer momento sobre el mismo software de Microsoft, para actualizar la hoja de datos a una versión más actual que permita tratar los datos en otro software. En este caso, el software usado para tratar los datos es el software *RStudio*. Este se basa en el lenguaje de programación R, que es una implementación de software libre del lenguaje S, desarrollado en 1993 por *R Development Core Team*, y descargable en <http://www.r-project.org>.

Finalmente, la metodología empleada se divide en dos partes:

- La primera, el uso de técnicas descriptivas. Estos son el cálculo de estadísticos descriptivos, el análisis gráfico de los datos y el uso de índices para medir la desigualdad en las distintas variables.
- La segunda, el uso de modelos de regresión. En este caso, modelos de regresión con datos de panel que permiten estudiar el comportamiento de varios sujetos a lo largo de un mismo período temporal.

## II. SITUACIÓN ACTUAL DEL MERCADO DE LA VIVIENDA

Uno de los principales problemas que ha dejado al descubierto la actual crisis económica en España es la dificultad de acceso a la vivienda. En otras palabras, mucha gente en este país tiene dificultades para poder acceder a uno de los derechos básicos. A través del mercado de la vivienda, el parque de la vivienda y la ocupación del suelo, este trabajo explicará la situación actual de España en referencia al mercado de la vivienda, los problemas que existen, analizará las acciones y reformas que se han llevado a cabo para garantizar este derecho de forma fácil a la población y propondrá algunas acciones que mejoren la situación actual.

El período de estudio comprende los años 2004 a 2014, once años que incluyen los tres previos a la actual crisis económica hasta el momento por el cual se nos ofrecen datos.

### 2.1. Situación actual en España

España, desde los años 60, ha aprovechado el sector de la construcción como importante para el desarrollo económico, dado que es un sector que genera empleo. Ligado con esta característica, desde la década de los 60, la política de vivienda se ha centrado en la tendencia a la propiedad por dos razones principales: la primera, esta política económica disminuye el gasto público; la segunda razón es la estabilidad social que genera este régimen de tenencia, ya que por razones sociológicas, la sociedad española ha preferido la estabilidad a través del ahorro frente al riesgo que representa la flexibilidad en otras inversiones más líquidas. Esta política, pues, transformó un país con un porcentaje de propietarios que vivían en sus propias viviendas de alrededor del 50% en la década de 1950 a un porcentaje superior al 80% en la década de 1980 (Rodríguez. R, 2010), siendo el más elevado de los países europeos.

**Tabla 2.1:** Número de viviendas construidas según año de construcción.

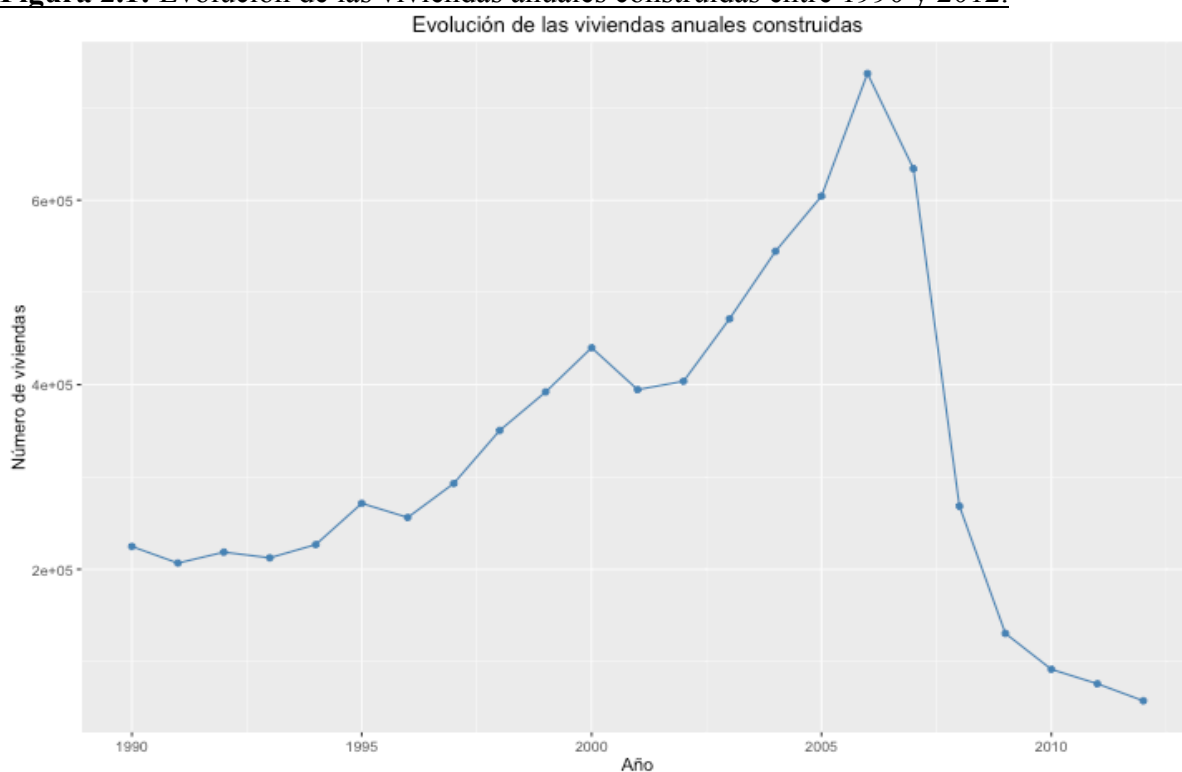
<i>Año de construcción</i>	<i>Número de viviendas</i>
1900-1920	590 700
1921-1940	754 815
1941-1950	830 840
1951-1960	1 834 105
1961-1970	3 573 200
1971-1980	5 011 355
1981-1990	3 137 195
1991-2001	3 396 195
2002-2011	4 349 420

Fuente: Ministerio de Fomento

En la Tabla 2.1 se puede observar como la construcción de nuevas viviendas ha sido muy importante en las últimas décadas. Destaca el crecimiento en la década de 1970, donde el sector inmobiliario creció y terminó con el estallido una burbuja inmobiliaria<sup>1</sup> en varios países, entre ellos España.

Si uno se fija en la evolución temporal de las dos últimas décadas (ver Figura 2.1), se observa un crecimiento del número de viviendas nuevas importante a partir de 1996, con un ligero retroceso en 2001 para, después de este año, crecer a un ritmo aún mayor hasta el año 2006, año previo al “pinchazo” de la burbuja inmobiliaria del año 2007. A partir de ese momento, en tres años se bajó al nivel de construcción que había tras la última burbuja inmobiliaria, que había estallado en 1992.

**Figura 2.1:** Evolución de las viviendas anuales construidas entre 1990 y 2012.



Fuente: Ministerio de Fomento. Elaboración propia.

Otra forma de ver la influencia del sector de la construcción en la economía es observar su tasa de paro. Esta, en 1990, partía de una situación de alrededor del 15% de desocupados en el sector de la construcción, aumentando hasta prácticamente el 30% debido al estallido de la burbuja de 1992. A partir de este momento, la tasa de paro disminuyó hasta el 5.99% del año 2006, pero en 3 años volvió a llegar a niveles superiores al 25%.

<sup>1</sup> AHEARNE, ALAN G.; AMMER, JOHN; DOYLE, BRIAN M.; KOLE, LINDA S.; MARTIN, ROBERT F. 2005. *International Finance Discussion Papers, Number 841. Board of Governors of the Federal Reserve System.*

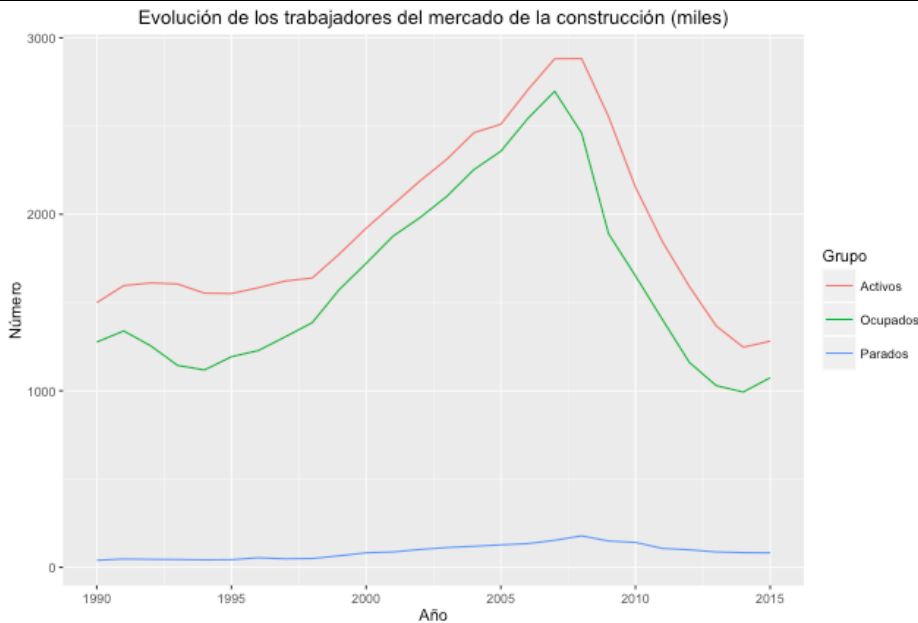
**Figura 2.2:** Evolución de la tasa de paro del sector de la construcción entre 1990 y 2015.



Fuente: Ministerio de Fomento. Elaboración propia.

Con la evolución de los activos, parados y ocupados del sector de la construcción, se puede observar el efecto sobre el mercado del trabajo que tiene este sector, con un fuerte incremento de los activos y los ocupados proporcional al crecimiento del sector. Por otra parte el número de parados se mantiene estable, y es la variación en la demanda del sector lo que hace variar la oferta, es decir, el crecimiento o decrecimiento del número de activos y ocupados y, en consecuencia, la tasa de paro del sector.

**Figura 2.3:** Evolución de los trabajadores del mercado de la construcción entre 1990 y 2015.



Fuente: Ministerio de Fomento. Elaboración propia.

Otra característica importante del sector de la construcción es efecto *bandwagon* o arrastre, el cual significa que a mayor demanda, mayor preferencia por un producto, y, por lo tanto,



lleva a un desequilibrio entre la oferta y la demanda. Esto ha llevado a la creación de un parque de vivienda sobredimensionado, con un exceso de oferta, que no ha sido cubierto ni por la demanda de segundas residencias ni por servicios como los destinados al turismo y que, además, ha sido sobrevalorado.

El valor de la vivienda depende del precio del suelo, de los costes de materiales y mano de obra y, finalmente, del margen que decida el promotor. El primero de los tres componentes, el precio del suelo, está regulado por el sector público. Esta regulación se tratará más adelante en este mismo capítulo.

### 2.1.1. Mercado de la vivienda

El mercado de la vivienda se puede dividir en tres mercados sustitutos: el mercado de la vivienda en propiedad, el mercado de la vivienda en alquiler y el mercado de la vivienda social.

#### 2.1.1.1. Mercado de la vivienda de propiedad

El mercado de vivienda español está formado, como se ha explicado anteriormente, por un porcentaje superior al 80% del mercado de vivienda en propiedad.

**Tabla 2.2:** Porcentaje de propietarios que viven en sus propias viviendas en España por comunidad autónoma en el 2º trimestre de 2014.

<i>Comunidad Autónoma</i>	<i>Porcentaje de propietarios dentro del total</i>
<i>Andalucía</i>	81.66
<i>Aragón</i>	79.62
<i>Principado de Asturias</i>	79.92
<i>Islas Baleares</i>	70.40
<i>Canarias</i>	70.77
<i>Cantabria</i>	81.46
<i>Castilla – La Mancha</i>	81.86
<i>Castilla y León</i>	81.74
<i>Cataluña</i>	74.32
<i>Comunidad Valenciana</i>	82.69
<i>Extremadura</i>	81.66
<i>Galicia</i>	77.86
<i>Comunidad de Madrid</i>	76.84
<i>Región de Murcia</i>	80.82
<i>Comunidad Foral de Navarra</i>	82.63
<i>País Vasco</i>	84.24
<i>La Rioja</i>	81.99
<i>Ceuta</i>	67.05
<i>Melilla</i>	66.97

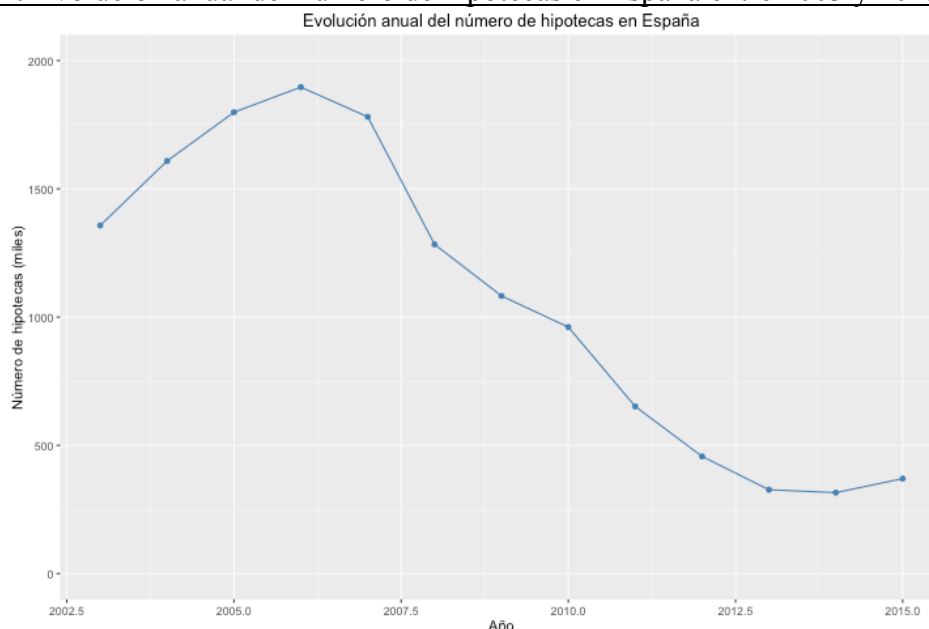
Fuente: Ministerio de Fomento. Elaboración propia.

El régimen de tenencia en España, destaca por la cantidad de propietarios que viven en sus propias viviendas, que se encuentra entre el 70% y el 85% sobre el total en todas las comunidades autónomas, con la excepción de las ciudades de Ceuta y Melilla, donde el porcentaje de propietarios que viven en sus viviendas es del 67%.

Otro hecho destacable, es la cantidad de propietarios que tienen una deuda sobre su propiedad.

Si se realiza un estudio histórico, partiendo de la década de 1980, donde el mercado de la propiedad representaba un porcentaje superior al 80% del acceso a la vivienda en España, y conociendo que esta característica se ha mantenido hasta los años previos a la crisis actual, se puede estudiar cómo se ha mantenido este mercado de la vivienda en propiedad y, sobretodo, cómo se ha financiado. Si uno se centra en los años para los cuales se tienen datos, el número de hipotecas firmadas aumentó hasta el año 2006, llegando a la cifra de casi 1 900 000 hipotecas, cifra muy superior a las 737 186 viviendas que se construyeron el mismo año en España y observando que estas dos variables siguen unas tendencias de crecimiento y decrecimiento parecidas para los años observados.

**Figura 2.4:** Evolución anual del número de hipotecas en España entre 2003 y 2015.



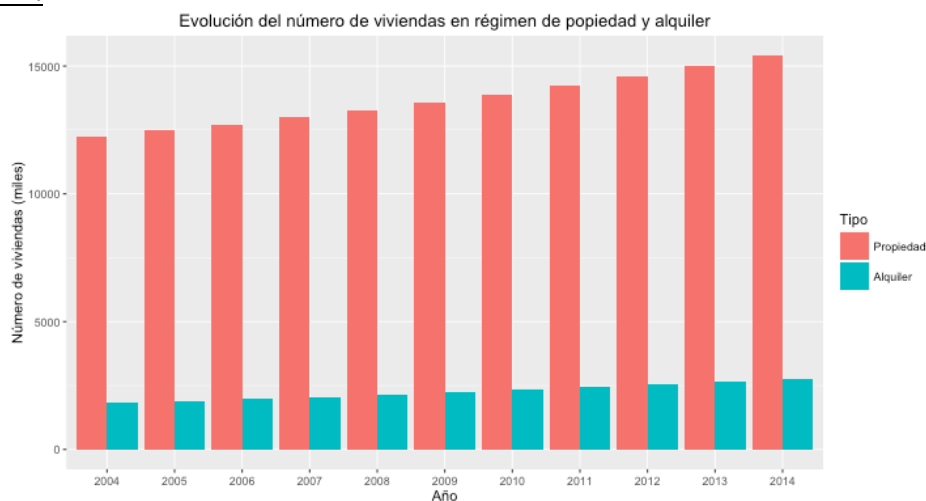
Fuente: INE. Elaboración propia.

La gran cantidad de hipotecas firmadas se debe a la cultura de la sociedad que ve en la propiedad una forma de ahorro y estabilidad frente a la situación inestable de vivir de alquiler. Además, a partir de 1996 los tipos de interés empezaron a disminuir como

consecuencia de la entrada de España en la Unión Monetaria Europea, haciendo la financiación de hipotecas barata en comparación con décadas anteriores. Cabe destacar que en España la cultura financiera es muy baja, con lo que se han firmado muchas hipotecas sin valorar la opción del alquiler o sin entender los efectos variables que pueden afectar a un préstamo hipotecario. Las consecuencias de esta forma de acceso a la vivienda se tratarán más adelante.

#### 2.1.1.2. Mercado de la vivienda de alquiler

**Figura 2.5:** Evolución del número de viviendas en régimen de propiedad y alquiler entre 2004 y 2014.

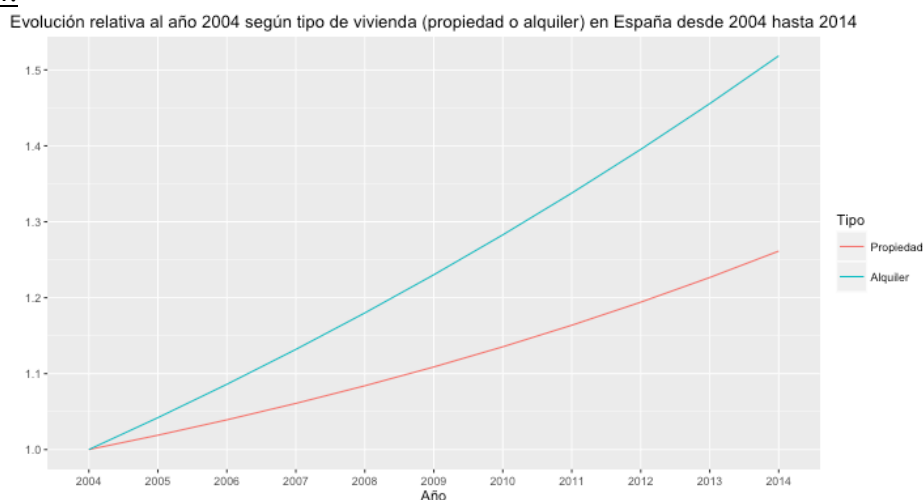


Fuente: Ministerio de Fomento. Elaboración propia.

En España, menos del 20% de las viviendas familiares principales del parque de la vivienda está destinado al alquiler. Este porcentaje, sin embargo, ha aumentado durante los últimos años, sobretodo durante la crisis, dado que muchas personas que no pueden permitirse financiar una vivienda en propiedad si pueden permitirse pagar un alquiler; por lo tanto, la demanda de vivienda en alquiler está aumentando. En la Figura 2.5, en la que sólo se han tenido en cuenta las viviendas en propiedad y alquiler del parque de vivienda español, se observa una evolución dado que en el año 2004 la propiedad representaba un 87.09% sobre el total de viviendas de propiedad y alquiler (12 237 730 viviendas en propiedad y 1 814 792 viviendas en alquiler), mientras que este porcentaje en el año 2014 era de 84.85% (15 436 555 viviendas en propiedad y 2 755 924 viviendas en alquiler), una reducción de un poco más de 2 puntos.

Otra forma de observar estos datos es mediante el crecimiento de ambas variables con base al año 2004.

**Figura 2.6:** Evolución relativa al año 2004 según el tipo de vivienda en España desde 2004 hasta 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

En la Figura 2.6, se observa que las viviendas principales en régimen de propiedad han crecido, desde 2004 hasta 2014, poco más del 25%, mientras que las viviendas principales destinadas al alquiler han crecido más del 50%.

Por el contrario, la oferta de la vivienda en alquiler no ha variado tanto ya que la rentabilidad obtenida por la venta de un inmueble dados los precios de los últimos años es mayor a la rentabilidad obtenida a través del alquiler de la vivienda. Sin embargo, la disminución de la demanda de compra de vivienda frente al aumento de la demanda de vivienda en alquiler ha llevado a propietarios a cambiar la naturaleza de su inversión y han puesto su inmueble en alquiler, aunque sea de forma temporal hasta poder vender.

Aun conociendo esta situación, el mercado del alquiler es un mercado regulado por el sector público a través de la Ley de Arrendamientos Urbanos, la cual durante las dos últimas décadas se ha utilizado para fomentar esta forma de acceso a la vivienda de la población española ya que, previo al año 1994, había básicamente dos tipos de alquileres:

1. Los basados en la Ley 40/1964, con rentas muy bajas.
2. Los basados en el Real Decreto-ley 2/1985, con rentas altas y de duración mayoritariamente anual.

En el primero de estos tipos de arrendamiento, el arrendador consideraba que estaba “perdiendo dinero”, mientras que en el segundo el que tenía esta sensación era el arrendatario, que veía en la propiedad una mejor forma de invertir su capital.

Con la Ley 29/1994 de Arrendamientos Urbanos, se quiso conseguir un punto de conciliación entre arrendador y arrendatario. La primera medida que se aplica con esta ley es la de duración mínima del contrato de arrendamiento de 5 años, junto con prorrogas de 3 años, para conseguir cierta estabilidad en ambas partes. En este período temporal se

considera que no es un plazo excesivo para que el propietario sitúe su vivienda en el mercado. Por otra parte, respecto a la renta, esta ley establece la libertad para pactar la renta inicial, pero determina que esta deba actualizarse con el Índice de Precios al Consumo. Finalmente, esta ley permite inscribir el contrato de arrendamiento en el Registro de la Propiedad, con los beneficios que esto pudiera suponer para ambas partes y también para el gobierno, dado que ello le permite obtener más datos sobre la evolución del mercado del alquiler.

La última gran modificación de la Ley de Arrendamientos Urbanos de 1994, la Ley 4/2013, pone en evidencia la necesidad de un mercado del alquiler fuerte, que en ese momento aún no existe en España, modificando la duración mínima del contrato reduciéndola de 5 a 3 años, y las prorrogas de 3 a 1 año. Además, dispone facilidades para que ambas partes puedan desistir del contrato. Por lo tanto, esta nueva ley lo que pretende es flexibilizar más el mercado.

Finalmente, con la Ley 2/2015, se abandona la indexación para la actualización de la renta del arrendamiento por una revisión periódica o no, no predeterminada, es decir, que ambas partes, arrendador y arrendatario, pacten la actualización de la renta. En caso de no acuerdo, la renta se actualiza anualmente a través del Índice de Garantía de Competitividad, que se define como:

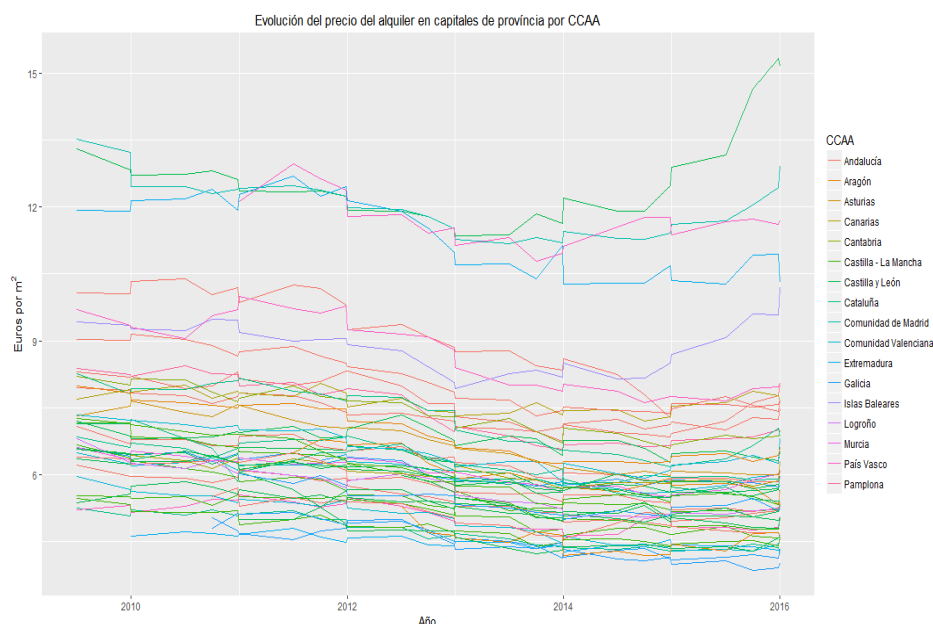
$$TV IGC_t = TV IPCA UEM_t - \alpha(TV IPCA ESP_{t,1999} - TV IPCA UEM_{t,1999}),$$

donde los datos se pueden obtener del *Eurostat* y representan:

- $TV IGC_t$  es la tasa de variación interanual del IGC en el mes  $t$ .
- $TV IPCA UEM_t$  es la tasa de variación interanual del IPC Armonizado en la Zona Euro en el mes  $t$ .
- $TV IPCA UEM_{t,1999}$  es la tasa de variación del IPC Armonizado de la Zona Euro entre el mes  $t$  y el mismo mes de 1999.
- $TV IPCA ESP_{t,1999}$  es la tasa de variación del IPC Armonizado de España entre el mes  $t$  y el mismo mes de 1999.

No hay estadísticas oficiales sobre el precio exacto de la vivienda en alquiler. Actualmente sólo existe una estadística del INE sobre la actualización del precio del alquiler. Sin embargo, grandes empresas del sector inmobiliario elaboran sus propias estadísticas que sirven como referencia para conocer el precio del alquiler y su evolución en España.

**Figura 2.7:** Evolución del precio del alquiler en capitales de provincia por CCAA.



Fuente: *Idealista.com*. Elaboración propia.

Para simplificar, en la Figura 2.7, se han representado las capitales de provincia bajo el nombre de la comunidad autónoma. En este caso se puede observar el comportamiento decreciente del precio por metro cuadrado hasta el año 2014. A partir de este año parece que empieza a crecer, donde destacan las cuatro curvas superiores. La que en la actualidad marca un nivel de precios más elevado es la provincia de Barcelona, seguida de la de Madrid, San Sebastián y Bilbao. Destaca también el crecimiento de los precios del alquiler de los dos últimos años en las Islas Baleares.

### 2.1.1.3. Mercado de la vivienda social

En España, el parque de la vivienda social no ha superado nunca el 2% del total del parque de viviendas principales<sup>2</sup>. Este dato es preocupante cuando se analiza el acceso a la vivienda, más en una situación de crisis económica donde el número de personas que no pueden permitirse comprar una vivienda o vivir de alquiler aumenta con el paso del tiempo y, sin embargo, no lo hace el parque de la vivienda social.

Este mercado, al igual que el del alquiler, está regulado, hecho que preocupa más la situación planteada en el párrafo anterior. La primera regulación en democracia de la vivienda social, llamada Vivienda de Protección Oficial (VPO), se realizó en 1978 con el Real Decreto-ley 31/1978 sobre políticas de viviendas de protección oficial. Este decreto ley estableció una única categoría de viviendas de protección oficial, que cumplieran las características de ser domicilio habitual, una superficie inferior o igual a 90 metros

<sup>2</sup> PITTINI, ALICE; LAINO, ELSA. 2011. Housing Europe Review 2012. The nuts and bolt of European social housing systems. Chapter II. CECODHAS Housing Europe's Observatory.

cuadrados y que cumplieran unos ciertos requisitos de precio y calidad, además de ser calificadas como VPO. También establecía que los promotores de la vivienda podían ser personas jurídicas o físicas, públicas o privadas y que los usuarios de las viviendas sólo podían ser personas físicas.

El régimen de estas viviendas podía ser arrendamiento o propiedad, y se establecía una ayuda a la financiación de la vivienda en función de las características de esa y de la renta de los usuarios.

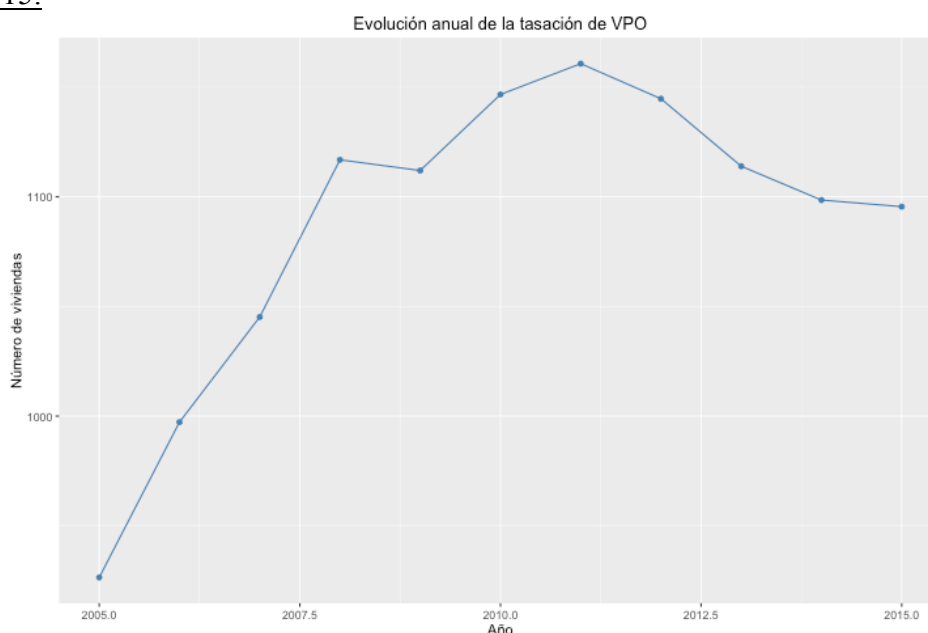
Con el plan de Vivienda 2009-2012, se actualiza la Ley 31/1978, estableciendo el mínimo de 30 metros cuadrados de superficie de la vivienda más 15 metros cuadrados por cada persona que habite en la vivienda. Además, se crean tres perfiles de VPO:

- Régimen especial, para ingresos máximos de 2.5 x IPREM
- Régimen general, para ingresos máximos de 4.5 x IPREM
- Régimen concertado, para ingresos máximos de 6.5 x IPREM

Donde el IPREM es el Indicador Público de Renta de Efectos Múltiples, el índice de referencia en España para la asignación de ayudas y subsidios en función de los ingresos.

Además, este plan establecía la duración mínima de arrendamiento de 10 o 25 años, 10 años para los arrendamientos con opción a compra y, finalmente, los préstamos para compra de VPO no podían ser inferiores a 10 o 25 años.

**Figura 2.8:** Evolución anual de la tasación de vivienda protegida oficial en España entre 2005 y 2015.



Fuente: Ministerio de Fomento. Elaboración propia.

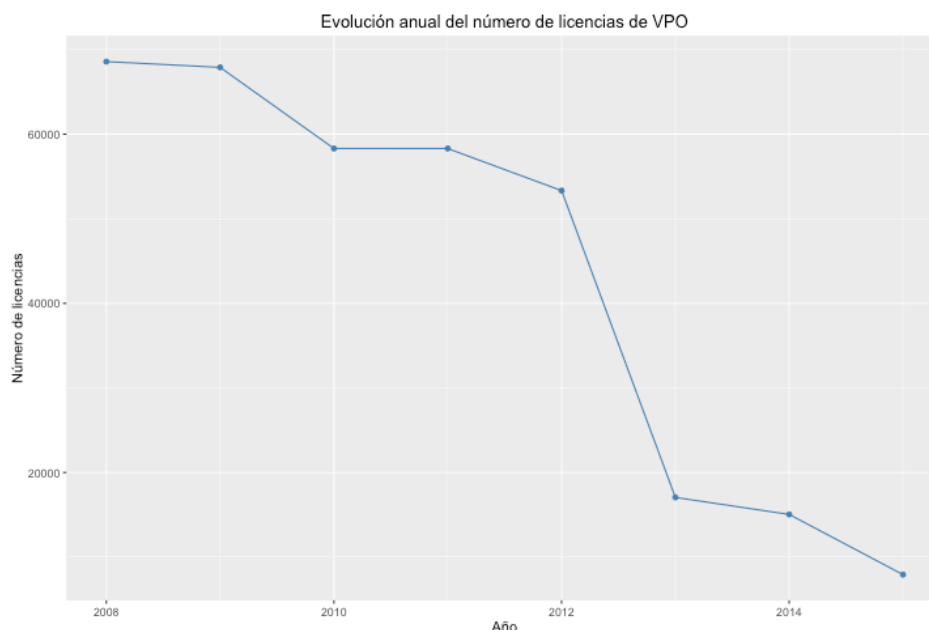
El valor del precio de la VPO está regulado por el gobierno, quien establece una rentabilidad máxima que se puede obtener con este tipo de viviendas. Además, su tasación depende del Módulo Básico Estatal (MBE), que es la cuantía en euros por metro cuadrado de superficie útil que sirve como referencia para la determinación de los precios máximos de venta. Según el MBE, los precios de la VPO son:

- Régimen especial, 1,5 veces el MBE.
- Régimen general, 1,6 veces el MBE.
- Régimen concertado, 1,8 veces el MBE.

Como se puede observar en la Figura 2.8, el precio de la VPO se ha estabilizado a partir de 2009.

Por otra parte, el número de licencias de VPO ha disminuido en los últimos años (ver Figura 2.9).

**Figura 2.9:** Evolución anual del número de licencias de vivienda protegida oficial entre 2008 y 2014.

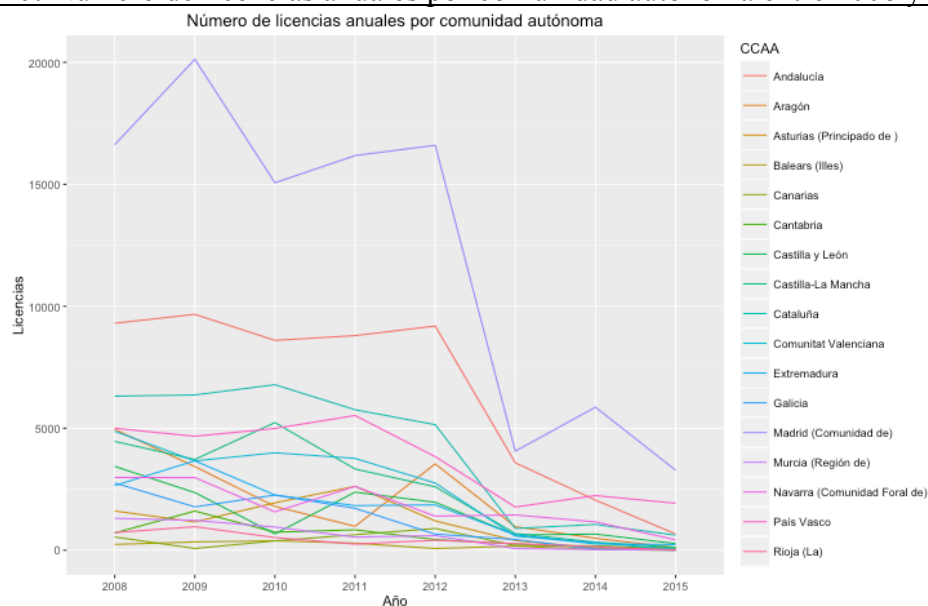


Fuente: Ministerio de Fomento. Elaboración propia.

A nivel autonómico, se puede observar también esta decaída, junto con el importante peso que tiene la Comunidad de Madrid en este mercado.



**Figura 2.10:** Número de licencias anuales por comunidad autónoma entre 2008 y 2015.



Fuente: Ministerio de Fomento. Elaboración propia.

### 2.1.2. Coste de la vivienda

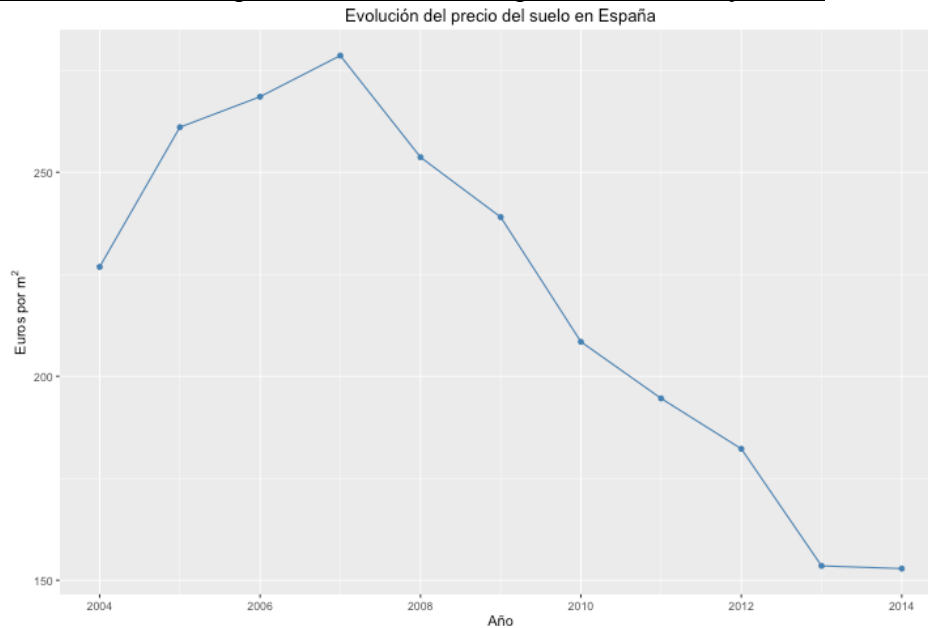
El coste de la vivienda se puede dividir en tres partes: precio del suelo, coste de la construcción y beneficio del promotor.

El primero de los tres factores, el precio del suelo, está regulado por el sector público. En esta regulación, se establecen cuatro elementos:

- El régimen urbanístico, que define y clasifica el tipo de suelo.
- El planeamiento urbanístico, que clasifica el suelo de un área concreta.
- La valoración de los tipos de suelo.
- La gestión urbanística, es decir, los procesos administrativos para la promoción y la construcción.

Existe una ley estatal sobre el suelo, la Ley de Suelo y Rehabilitación Urbana. Esta define el régimen urbanístico. Sin embargo, en España cada comunidad autónoma tiene su propia ley sobre la materia, y son los gobiernos de estas las que aprueban los planes urbanísticos de cada municipio. Los planes urbanísticos son definidos por las administraciones locales.

**Figura 2.11:** Evolución del precio del suelo en España entre 2004 y 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

Por lo tanto, encontramos una oferta de suelo regulada en cuanto a cantidad. El precio del suelo depende del equilibrio entre la oferta y la demanda, por lo tanto, a más demanda, mayor precio del suelo. Como se puede observar en el gráfico anterior, la fuerte demanda de la burbuja inmobiliaria hizo que el precio del suelo aumentara. Con el pinchazo de la burbuja y la caída de la demanda, el precio empezó a caer.

El segundo de los factores, el coste de la construcción, formado por el coste de los materiales y el coste de la mano de obra, funciona aproximadamente como un mercado de competencia perfecta. El coste de la mano de obra creció mientras crecía la burbuja dado el exceso de demanda.

El tercer de los factores es el margen de beneficio del promotor. Este no tiene regulación.

## 2.2. Problemas del mercado de la vivienda español

Las ventajas que ofrece un mercado basado en el acceso a la vivienda a través de la propiedad, por el contrario, quedan descompensadas por los inconvenientes que genera ese tipo de acceso. El principal inconveniente es la dificultad del acceso a la vivienda por parte de las personas con menos renta de una población. Estas personas no tienen forma de financiarse una vivienda. Además, algunas de ellas también tienen dificultades para poder pagar un alquiler. Ligado con este problema, a lo largo de los últimos años el número de personas con rentas bajas en el Estado español ha aumentado. Aquí se presenta la necesidad de políticas económicas asignativas para mejorar la oferta del mercado de la

vivienda y políticas económicas redistributivas para mejorar la demanda del mercado de la vivienda a través del acceso a esta.

Otra característica negativa de la propiedad como acceso a la vivienda es la drástica reducción de la movilidad como consecuencia, es decir, las dificultades para cambiar de residencia ya sea para la búsqueda de empleo o para un cambio de residencia con otras finalidades, como la jubilación en un territorio con un clima más suave que el del sitio de procedencia.

### III. DATOS DE ESTUDIO

#### 3.1. Fuente de datos. Atlas Digital

Para este trabajo, los datos de estudio serán los de las base de datos del Atlas Digital. El Atlas Digital es un geoportal del Ministerio de Fomento que, mediante una interfaz sencilla e intuitiva, permite al usuario elaborar cartografía temática combinando diferentes variables a diferentes niveles geográficos: Comunidades Autónomas, Provincias, Áreas Urbanas y Municipios.

Para tener un estudio más completo, se pretendía trabajar a nivel municipal, pero dado el elevado nivel de datos *missing*, superior al 95%, se ha decidido hacer el estudio sobre las variables de las áreas urbanas y las comunidades autónomas.

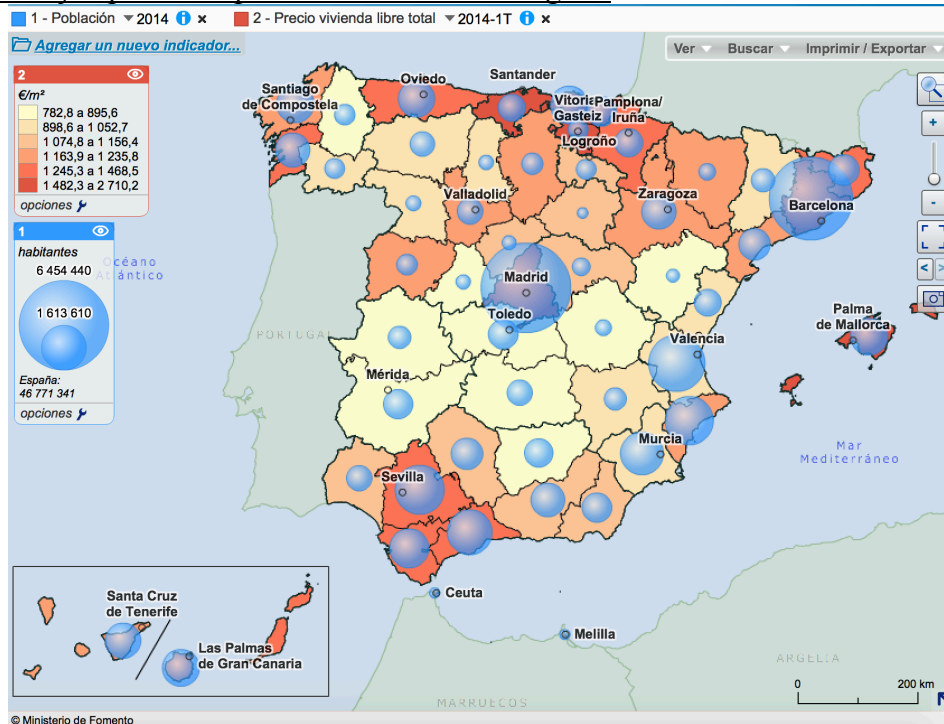
Los datos que se obtienen del Atlas Digital se extraen con un archivo Excel de versión anterior al año 95, hecho que obliga a tratar los datos en la misma hoja de cálculo ya que estos no pueden ser leídos por el software R. Además, hay que tratar, también en la hoja de cálculo, los datos *missing*, para que estos puedan ser leídos como tal en R.

##### 3.1.1. Funcionamiento del Atlas Digital

Del Atlas Digital se pueden obtener “Mapas temáticos”, “Informes” y “Tablas de datos”, a los cuales se puede acceder a través de los menús de la parte superior derecha de la pantalla. En los primeros, se pueden representar hasta dos variables de las del Atlas Digital en un mapa según el nivel geográfico elegido: comunidad autónoma, provincia, área urbana o municipio, según disponibilidad de los datos.

En el ejemplo de la Figura 3.1 se representan sobre el mapa de las provincias españolas las variables de “Población” en el año 2014 y “Precio de la vivienda nueva” en el primer trimestre de 2014. El Atlas Digital, permite ver dos mapas a la vez, así como también seleccionar las regiones que interese estudiar o obtener datos de ellas. También permite filtrar los datos, como por ejemplo escoger las poblaciones de más de cincuenta mil habitantes.

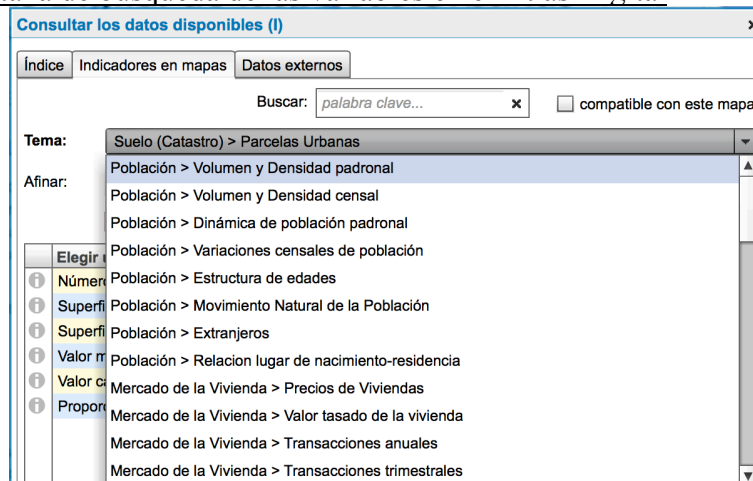
**Figura 3.1: Ejemplo de mapa temático del Atlas Digital.**



Fuente: Ministerio de Fomento.

A partir de los datos seleccionados, se pueden crear automáticamente breves informes, como el que se encuentra de ejemplo en la página web<sup>3</sup> del Atlas Digital sobre la evolución de la población de la Comunidad de Madrid. Estos datos se eligen a través de la ventana de búsqueda de variables, donde estas están ordenadas por tema y se filtran según nivel geográfico. El período temporal se elige a posteriori, cuando las variables se proyectan en los mapas temáticos.

**Figura 3.2: Ventana de búsqueda de las variables en el Atlas Digital**



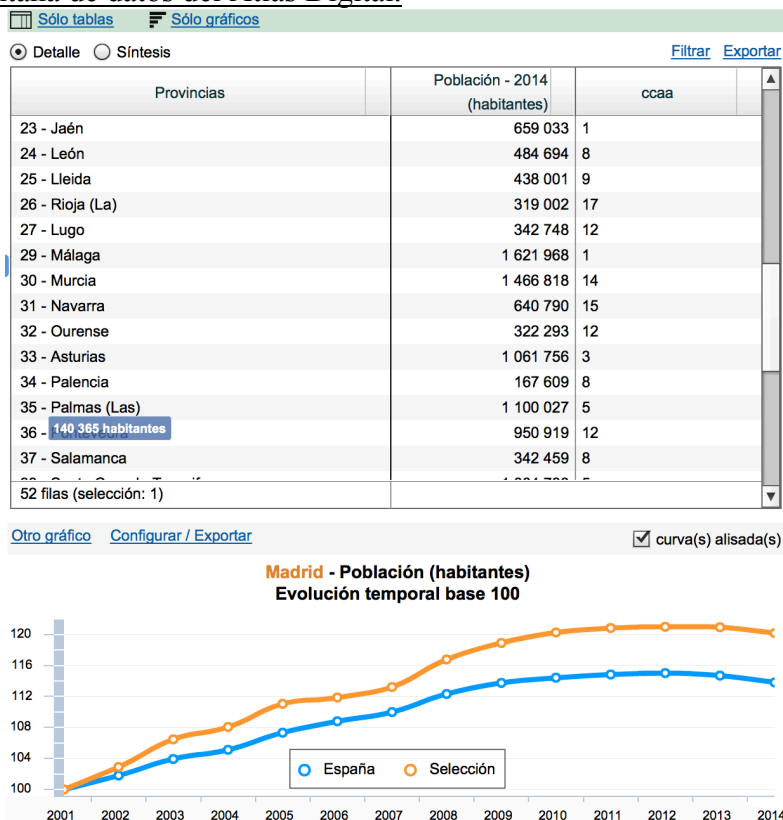
Fuente: Ministerio de Fomento.

3

[http://atlasau.fomento.gob.es/GC\\_preport.php?lang=es&nivgeo=prov&view=map2&codigo=28&id\\_rep=r01](http://atlasau.fomento.gob.es/GC_preport.php?lang=es&nivgeo=prov&view=map2&codigo=28&id_rep=r01)

Finalmente, en el apartado “Datos” encontramos las tablas con los datos seleccionados. Además, en este apartado se pueden representar gráficamente las variables seleccionadas. Si se quiere exportar los datos, en el apartado exportar, como se observa en la Figura 3.3, se permite exportar los datos en versión *pdf* o en Excel de formato citado anteriormente. Se pueden descargar las variables disponibles que se deseen, independientemente de su tipo. El período temporal elegido es también independiente por cada variable, pero no el nivel geográfico. Los datos deben descargarse para un mismo nivel geográfico.

**Figura 3.3:** Ventana de datos del Atlas Digital.



Fuente: Ministerio de Fomento.

### 3.2. Variables de estudio

Se toman variables según la siguiente clasificación: variables demográficas, variables del mercado de vivienda, variables del parque de vivienda y variables de la ocupación del suelo según la clasificación CORINE (de las siglas de *Coordination of Information on the Environment, Land Cover*), el cual es un proyecto con el objetivo de desarrollar la creación de una base de datos sobre la cobertura y uso del territorio dentro de la Unión Europea. Estas variables se pueden obtener según nivel geográfico y período. Además, se presenta la unidad en la que está expresada cada variable. Las definiciones se pueden encontrar en el Anexo.

### 3.2.1. Variables demográficas

**Tabla 3.1:** Variables demográficas del Atlas Digital tomadas para el estudio.

<i>Variable</i>	<i>Niveles geográficos</i>	<i>Periodos</i>	<i>Unidad</i>
<i>Población</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2001-2014	Número de habitantes
<i>Densidad de población</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2001-2014	Habitantes por km <sup>2</sup>
<i>Extranjeros</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2001-2014	Número de habitantes extranjeros
<i>Edad media de la población</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2001-2014	Número

Fuente: Ministerio de Fomento. Elaboración propia.

### 3.2.2. Variables del mercado de la vivienda

**Tabla 3.2:** Variables del mercado de la vivienda del Atlas Digital tomadas para el estudio.

<i>Variable</i>	<i>Niveles geográficos</i>	<i>Periodos</i>	<i>Unidad</i>
<i>Precio vivienda libre total</i>	Comunidad Autónoma, Provincia, Municipio	1995 - 2014	Euros por m <sup>2</sup>
<i>Precio vivienda libre nueva</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2004 - 2014	Euros por m <sup>2</sup>
<i>Precio vivienda libre usada</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	1995 - 2014	Euros por m <sup>2</sup>
<i>Valor tasado vivienda total</i>	Comunidad Autónoma, Provincia, Municipio	1995 - 2014	Euros por m <sup>2</sup>
<i>Valor tasado vivienda nueva</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2014 - 2015	Euros por m <sup>2</sup>
<i>Valor tasado vivienda usada</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2014 - 2015	Euros por m <sup>2</sup>
<i>Número de transacciones anuales</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2004 - 2014	Número de transacciones

Fuente: Ministerio de Fomento. Elaboración propia.

Los datos referentes al precio y la tasación hacen referencia al precio medio por cada región en el momento temporal dado.

### 3.2.3. Variables del parque de vivienda

#### 3.2.3.1. Variables referentes al tamaño del parque de vivienda

**Tabla 3.3:** Variables del parque de la vivienda referentes al tamaño del parque del Atlas Digital tomadas para el estudio.

<b>Variable</b>	<b>Niveles geográficos</b>	<b>Periodos</b>	<b>Unidad</b>
<i>Densidad de viviendas familiares</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2011, 2011	Viviendas por km <sup>2</sup>
<i>Densidad de viviendas familiares principales</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2001, 2011	Viviendas por km <sup>2</sup>
<i>Densidad de viviendas secundarias</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2001, 2011	Viviendas por km <sup>2</sup>
<i>Densidad de viviendas vacías</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2001, 2011	Viviendas por km <sup>2</sup>
<i>Viviendas familiares principales</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2001, 2011	Número de viviendas
<i>Viviendas secundarias</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2001, 2011	Número de viviendas
<i>Viviendas vacías</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2001, 2011	Número de viviendas
<i>Otras viviendas</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2001, 2011	Número de viviendas

Fuente: Ministerio de Fomento. Elaboración propia.

#### 3.2.3.2. Variables referentes al régimen de tenencia de las viviendas familiares principales

**Tabla 3.4:** Variables del parque de la vivienda referentes al régimen de tenencia de las viviendas familiares principales del Atlas Digital tomadas para el estudio.

<b>Régimen de tenencia</b>	<b>Niveles geográficos</b>	<b>Periodos</b>	<b>Unidad</b>
<i>Propiedad pagada</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2001	Número de viviendas
<i>Pagos pendientes</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2001	Número de viviendas
<i>Propiedad por herencia o donación</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2001	Número de viviendas
<i>Alquiler</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2001	Número de viviendas
<i>Régimen de tenencia gratuito</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2001	Número de viviendas
<i>Otro régimen de tenencia</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	2001	Número de viviendas

Fuente: Ministerio de Fomento. Elaboración propia.



### 3.2.4. Variables de la ocupación del suelo

**Tabla 3.5:** Variables de la ocupación del suelo del Atlas Digital tomadas para el estudio.

<i>Variable</i>	<i>Niveles geográficos</i>	<i>Periodos</i>	<i>Unidad</i>
<i>Superficie de suelo artificial</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	1987, 2000, 2006	Hectáreas
<i>Tejido urbano continuo</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	1987, 2000, 2006	Porcentaje
<i>Zonas industriales y comerciales</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	1987, 2000, 2006	Porcentaje
<i>Redes viarias y ferroviarias</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	1987, 2000, 2006	Porcentaje
<i>Zonas portuarias</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	1987, 2000, 2006	Porcentaje
<i>Aeropuertos</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	1987, 2000, 2006	Porcentaje
<i>Zona de extracción minera</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	1987, 2000, 2006	Porcentaje
<i>Escombreras y vertederos</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	1987, 2000, 2006	Porcentaje
<i>Zonas en construcción</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	1987, 2000, 2006	Porcentaje
<i>Zonas verdes urbanas</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	1987, 2000, 2006	Porcentaje
<i>Instalaciones deportivas y recreativas</i>	Comunidad Autónoma, Provincia, Área Urbana, Municipio	1987, 2000, 2006	Porcentaje

Fuente: Ministerio de Fomento. Elaboración propia.

### 3.3. Descriptiva de los datos

A través del análisis exploratorio de los datos, se hace un primer análisis de la información que se puede obtener de estos.

El primer análisis se lleva a cabo a nivel municipal. En España hay un total de 8198 municipios.

### 3.3.1. Descriptiva por municipios

#### 3.3.1.1. Descriptiva de las variables demográficas

**Tabla 3.6:** Descriptivos de las variables demográficas de los datos según municipio.

	<i>Mínimo</i>	<i>Q1</i>	<i>Q2</i>	<i>Media</i>	<i>Q3</i>	<i>Máximo</i>	<i>Missings</i>
<i>Población</i>	3	166	556	5762	2443	3165235	81
<i>Densidad de población</i>	0.2	5.2	14.2	175.7	55.7	22800	82
<i>Extranjeros</i>	0	5	25	618	147	417731	81
<i>Edad media</i>	31.2	42.5	47.3	47.7	52.4	69.4	81

Fuente: Ministerio de Fomento. Elaboración propia.

De los 8198 municipios, hay 81 de los cuales no se tienen datos.

Por su parte, viendo los datos de población, se puede comprobar como hay pocas ciudades con muchos habitantes, siendo el máximo Madrid con 3 165 235 habitantes en 2014. Se puede comprobar como el tercer cuartil de esta variable es de 2443 habitantes, es decir, el 75% de los municipios españoles tienen 2443 o menos habitantes, y el 50% tienen menos de 556 habitantes, cantidad marcada por la mediana.

Respecto a la densidad de población, la mayoría de municipios españoles son poco densos, donde el 75% de los municipios tienen una densidad igual o inferior a los 55.7 habitantes por kilómetro cuadrado. Aunque la densidad de la mayoría de municipios sea baja, encontramos municipios con alta densidad de población, siendo el máximo en 22800 habitantes por kilómetro cuadrado en el pueblo de Emperador, en la provincia de Valencia. Este municipio tiene, en 2014, 684 habitantes, pero cuenta con una superficie de 0.03 kilómetros cuadrados, hecho que provoca su alta densidad. Por otra parte, tenemos el ejemplo de una gran ciudad como Barcelona, con 1 602 386 habitantes y una densidad de población de 16315.9 habitantes por kilómetro cuadrado.

Por su parte, la población inmigrante está muy concentrada, siendo el máximo de población extranjera de 417 731 habitantes en un mismo municipio.

Finalmente, las edades medias de la población se encuentran entre los 31.2 años y los 69.4, pero el 50% de la población está comprendida entre los 42.5 y los 52.4 años, siendo la media poblacional de 47.7 años.

#### 3.3.1.2. Descriptiva de las variables del mercado de vivienda

**Tabla 3.7:** Descriptivos de las variables del mercado de la vivienda de los datos según municipio.

	<i>Mín.</i>	<i>Q1</i>	<i>Q2</i>	<i>Media</i>	<i>Q3</i>	<i>Máx.</i>	<i>Missing</i>
<i>Precio de la vivienda total</i>	544.7	912.2	1119.2	1226.4	1422.5	3298.1	7919
<i>Tasación de la vivienda total</i>	1.0	1.3	2.0	3656.1	8092.5	9998.0	7990
<i>Tasación de la vivienda nueva</i>	2551	11878	15305	14774	17391	23817	8166
<i>Tasación de la vivienda usada</i>	1.00	1.30	2.05	3765.78	8089.0	9984.00	7992

Fuente: Ministerio de Fomento. Elaboración propia

### 3.3.1.3. Descriptiva de las variables del parque de vivienda

**Tabla 3.8:** Descriptivos de las variables del parque de la vivienda de los datos según municipio.

	<b>Mín.</b>	<b>Q1</b>	<b>Q2</b>	<b>Media</b>	<b>Q3</b>	<b>Máx.</b>	<b>Missing</b>
<i>Densidad de viviendas familiares</i>	2.90	26.40	65.95	274.39	204.65	10584.50	5892
<i>Densidad de viviendas familiares principales</i>	1.90	17.60	44.15	209.45	141.90	8749.50	5892
<i>Densidad de viviendas secundarias</i>	0.00	2.90	7.30	31.39	19.88	1583.80	5892
<i>Densidad de viviendas vacías</i>	0.000	3.825	9.400	33.559	27.500	1.421.000	5892
<i>Viviendas familiares principales</i>	652	1225	2217	7344	5166	1321064	5891
<i>Viviendas secundarias</i>	0.0	186.5	421.0	1238.1	953.5	62585.0	5891
<i>Viviendas vacías</i>	0.00	257	502	1347	1126	153101	5891
<i>Otras viviendas</i>	0.000	0.000	1.000	3.626	3.000	533.000	5891
<i>Viviendas en propiedad pagada</i>	443	1916	2799	7364	5123	546787	7443
<i>Viviendas en propiedad con pagos pendientes</i>	257	1745	2938	6459	5475	369858	7443
<i>Propiedad por herencia o donación</i>	112	368	587	1032	929	49921	7489
<i>Alquiler</i>	118	501	965	2903	1995	264131	7467
<i>Gratuita</i>	94.0	212.0	289.0	640.1	522.0	31143.0	7753
<i>Otro régimen</i>	94.0	281.0	421.0	913.5	716.0	58691.0	7467

Fuente: Ministerio de Fomento. Elaboración propia.

Dada la alta cantidad de NA (superior al 96.5% de los datos), se desestima el trabajo con municipios y se decide tomar el área urbana como unidad mínima geográfica de trabajo en este estudio.

### 3.3.2. Descriptiva por Comunidades Autónomas

Para una mayor visualización de la situación actual que nos muestran los datos del Atlas Digital, se procede a hacer una descriptiva de las variables a través del menor número de sujetos posible, que es a través de las comunidades autónomas. En España hay un total de 17 comunidades autónomas.

Para hacer la descriptiva más clara, se procesan los siguientes cambios:

<b>CCAA</b>	<b>Abreviatura</b>
Andalucía	AND
Aragón	ARA
Asturias (Principado de)	AST
Baleares (Illes)	BAL
Canarias	CAN
Cantabria	CNT
Castilla - La Mancha	CLM
Castilla y León	CyL
Cataluña	CAT
Comunidad Valenciana	VAL
Extremadura	EXT
Galicia	GAL
Madrid (Comunidad de)	MAD
Murcia (Región de)	MUR
Navarra (Comunidad Foral de)	NAV
País Vasco	PVA
Rioja (La)	RIO

### 3.3.2.1. Descriptiva de las variables demográficas

Se presentan los principales estadísticos descriptivos (mínimo, primer cuartil, mediana, media, tercer cuartil y máximo) para las cuatro variables numéricas de carácter demográfico, que son la *población*, *densidad de población*, *extranjeros* y *edad media* con datos de 2014.

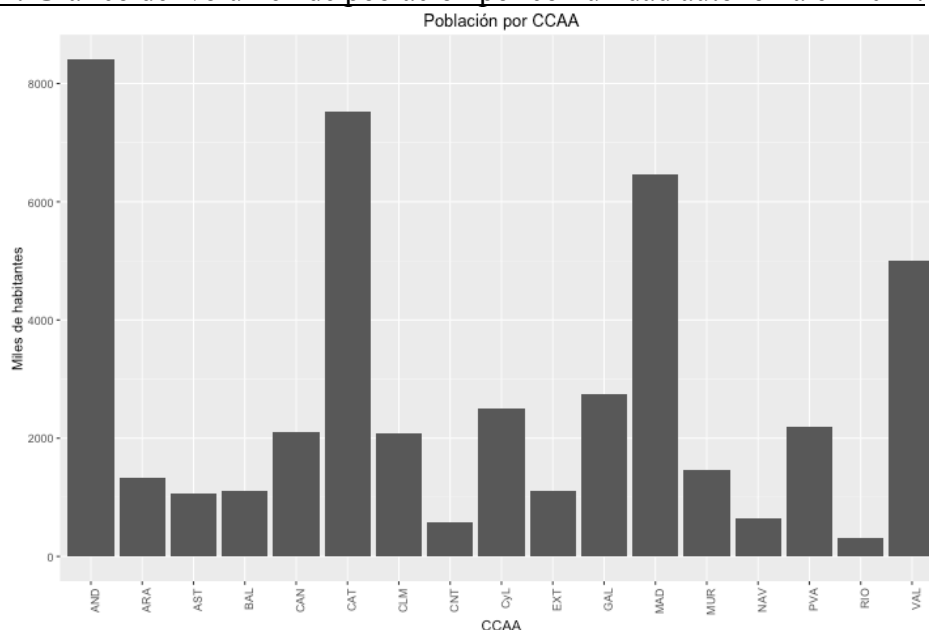
**Tabla 3.9:** Descriptivos de las variables demográficas de los datos según comunidad autónoma.

	<b>Mín.</b>	<b>Q1</b>	<b>Q2</b>	<b>Media</b>	<b>Q3</b>	<b>Máx.</b>
<i>Población</i>	319002	1099632	2078611	2741286	2748695	8402305
<i>Densidad de población</i>	26.2	63.4	100.1	166.6	221.1	804.6
<i>Extranjeros</i>	15027	59550	131927	209337	267138	879953
<i>Edad media</i>	38.80	40.80	41.90	42.24	43.60	45.90

Fuente: Ministerio de Fomento. Elaboración propia.

- Población: la comunidad autónoma con más habitantes es Andalucía, con 8 402 305 habitantes y la que menos La Rioja, con 319 002 habitantes. La media de habitantes por comunidad autónoma es de 2 741 286.

**Figura 3.4:** Gráfico del volumen de población por comunidad autónoma en 2014.



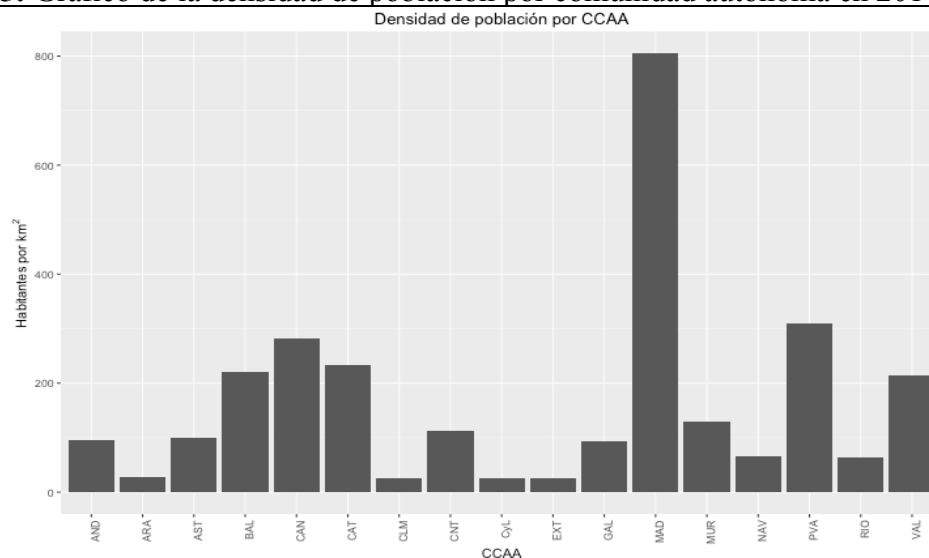
Fuente: Ministerio de Fomento. Elaboración propia.

- Densidad de población: la comunidad con más densidad de población, siendo ésta más del doble que la siguiente con más densidad, es la Comunidad de Madrid, con una densidad de 804.6 habitantes por kilómetro cuadrado. Las comunidades autónomas con menos densidad de población son Castilla - La Mancha, Extremadura y Castilla y León, con 26.2, 26.4 y 26.6 habitantes por kilómetro cuadrado respectivamente.

Con excepción de la Comunidad de Madrid, las comunidades con menos densidad de población son las de interior, mientras que las más densas son las de costa e islas.

La media de la densidad de población de las comunidades autónomas es de 166.6 habitantes por kilómetro cuadrado.

**Figura 3.5:** Gráfico de la densidad de población por comunidad autónoma en 2014.

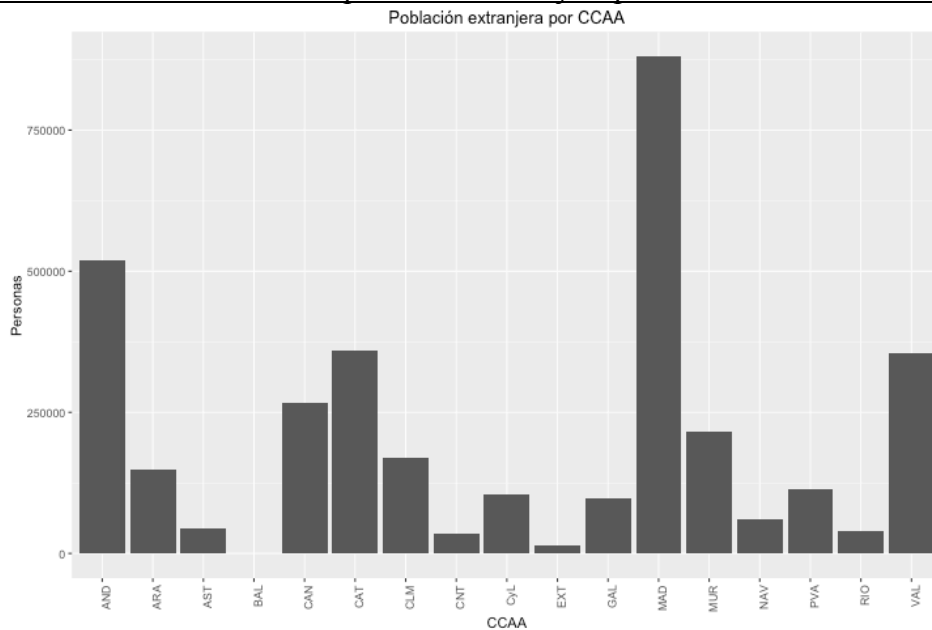


Fuente: Ministerio de Fomento. Elaboración propia.

- Población extranjera: la comunidad autónoma con más población extranjera es la Comunidad de Madrid, con una población extranjera de 879 952 personas, seguida por Andalucía, con más de medio millón de personas de procedencia extranjera. La región con menos población extranjera es Extremadura, con sólo 15 027 personas de procedencia extranjera. La media de personas extranjeras por comunidad autónoma es de 214 175 personas.

No hay datos de las Islas Baleares.

**Figura 3.6:** Gráfico del volumen de población extranjera por comunidad autónoma en 2014.

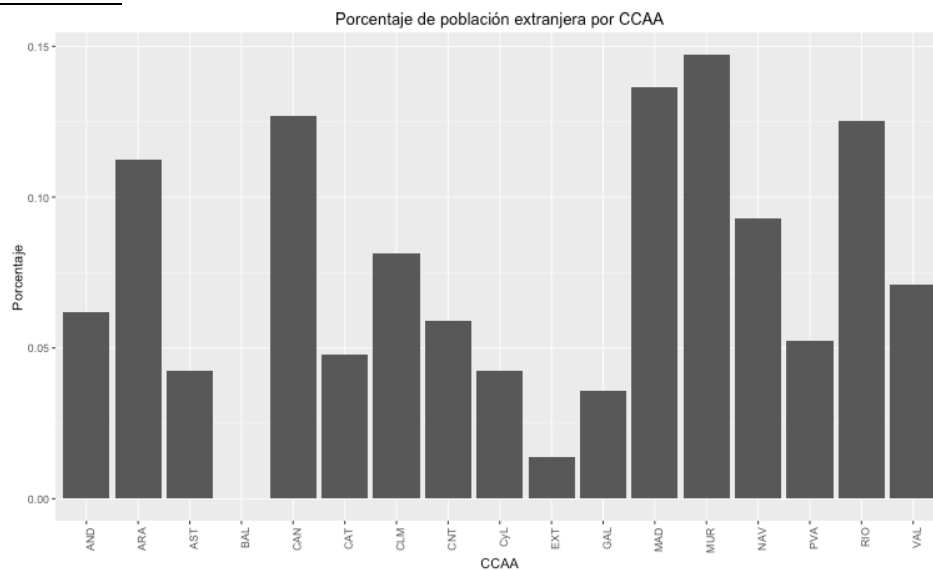


Fuente: Ministerio de Fomento. Elaboración propia.

Junto con los valores absolutos de población extranjera, se puede observar que porcentaje de la población extranjera sobre el total de la población de cada comunidad autónoma.

En este caso no se observa tanta dispersión como en el caso de los valores absolutos, pero si se observan regiones, como la Comunidad de Madrid o Murcia, donde el porcentaje de población extranjera respecto a la población total es superior al doble de la media del porcentaje de población extranjera sobre el total de la población por comunidad autónoma.

**Figura 3.7:** Gráfico de la población extranjera respecto a la población total por comunidad autónoma en 2014.

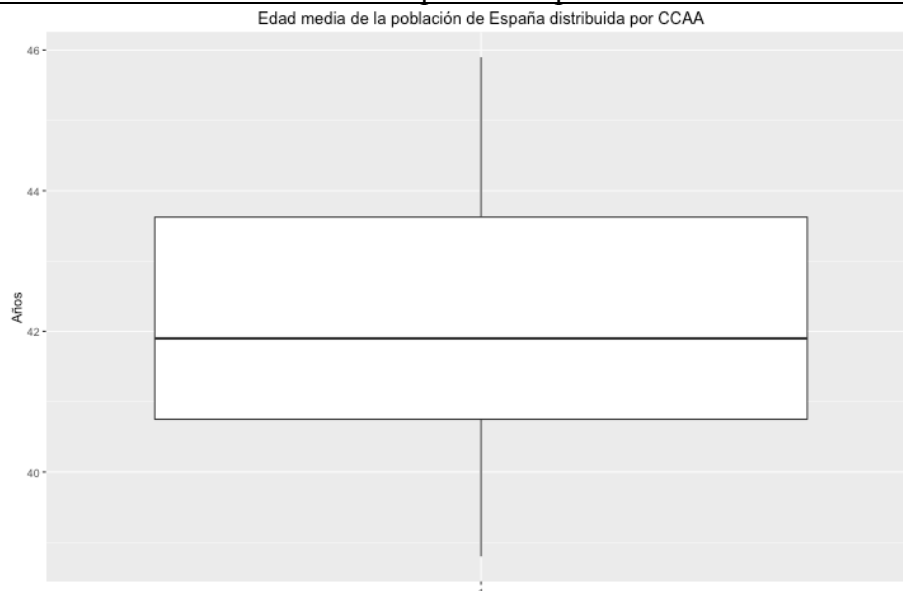


Fuente: Ministerio de Fomento. Elaboración propia.

- Edad media: la edad media de las comunidades autónomas se encuentra en 42.26 años. La comunidad autónoma más envejecida es Asturias, con una edad media de 45.9 años y la comunidad autónoma con la población más joven es Murcia, con una edad media 38.8 años.

No hay datos de las Islas Baleares.

**Figura 3.8:** Gráfico de la edad media de la población por comunidad autónoma en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

### 3.3.2.2. Descriptiva de las variables del mercado de la vivienda

Se presentan los principales estadísticos descriptivos (mínimo, primer cuartil, mediana, media, tercer cuartil y máximo) para las ocho variables numéricas del mercado de vivienda, que son “precio vivienda libre total”, “precio vivienda libre nueva”, “precio vivienda libre usada”, “tasación vivienda libre total”, “tasación vivienda libre nueva”, “tasación vivienda libre usada” y “transacciones anuales” con datos de 2014.

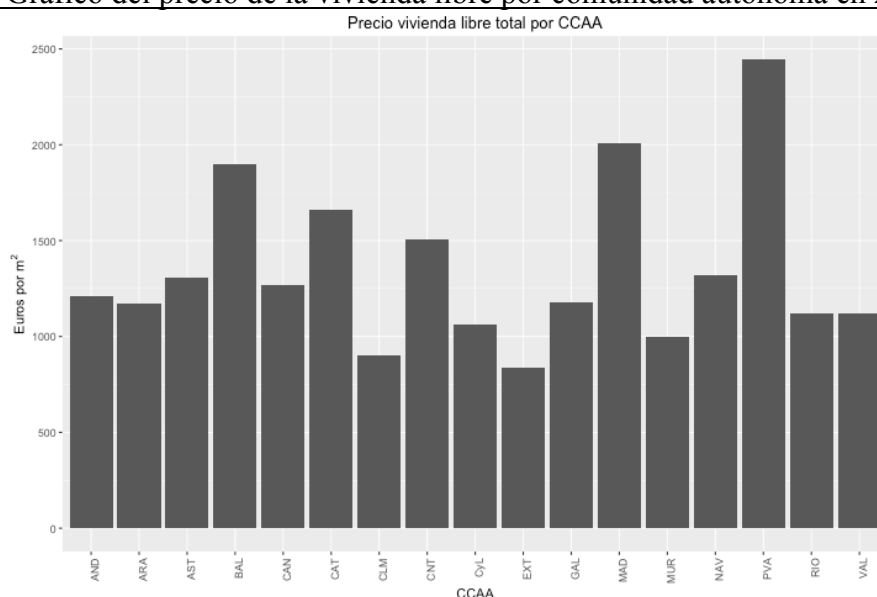
**Tabla 3.10:** Descriptivos de las variables del mercado de la vivienda de los datos según comunidad autónoma.

	<i>Mín.</i>	<i>Q1</i>	<i>Q2</i>	<i>Media</i>	<i>Q3</i>	<i>Máx.</i>
<i>Precio vivienda total</i>	837.1	1118.0	1207.1	1353.2	1504.7	2443.4
<i>Precio vivienda nueva</i>	941.1	1267.6	1316.2	1478.3	1661.3	2559.2
<i>Precio vivienda usada</i>	806.5	1018.0	1156.8	1292.9	1474.6	2385.4
<i>Tasación vivienda total</i>	837.1	1118.0	1207.1	1353.2	1504.7	2443.4
<i>Tasación vivienda nueva</i>	985.1	1386.4	1473.5	1609.6	1715.0	2736.1
<i>Tasación vivienda usada</i>	830.8	1094.9	1203.7	1342.1	1494.0	2430.2
<i>Transacciones anuales</i>	2753	6282	14009	21453	17915	69999

Fuente: Ministerio de Fomento. Elaboración propia.

- Precio vivienda libre total: las comunidades autónomas con precios más elevados de vivienda, por encima de la media de las comunidades autónomas que se encuentra en 1353.2 euros el metro cuadrado, son el País Vasco, la Comunidad de Madrid, las Islas Baleares y Cataluña, con 2443.4, 2008.1, 1901 y 1658.3 euros por metro cuadrado respectivamente. Las comunidades con los precios más baratos son Extremadura y Castilla - La Mancha, ambas con precios inferiores a los 1000 euros por metro cuadrado.

**Figura 3.9:** Gráfico del precio de la vivienda libre por comunidad autónoma en 2014.

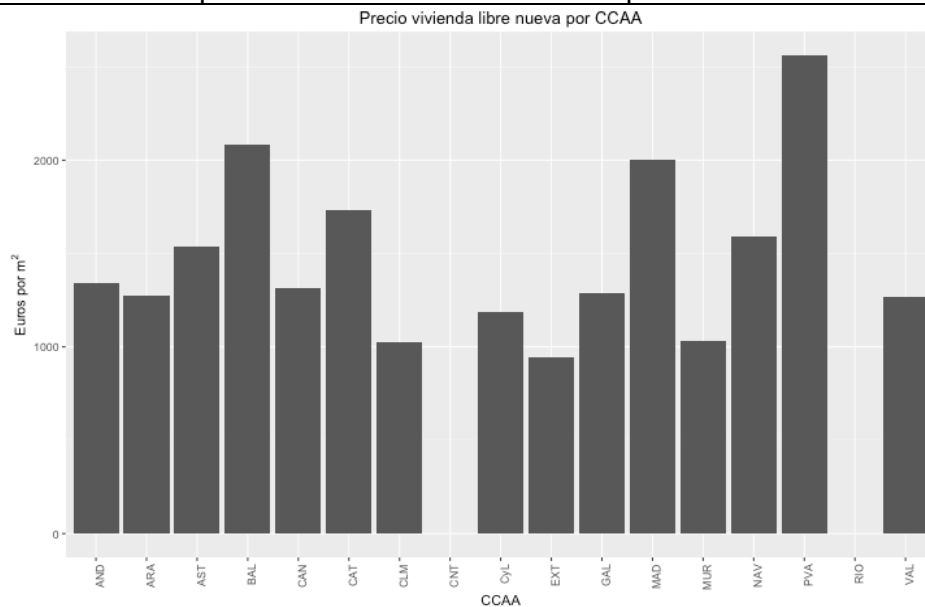


Fuente: Ministerio de Fomento. Elaboración propia.



- Precio vivienda libre nueva: los precios de la vivienda libre nueva son, en todos los casos menos en la Comunidad de Madrid, más elevados que los precios del total de la vivienda libre. Las comunidades con los precios más elevados siguen siendo las citadas en la variable anterior, siendo el máximo del País Vasco en 2559.2 euros el metro cuadrado y el mínimo en 941.1 euros el metro cuadrado en Extremadura. No hay datos de Cantabria y La Rioja.

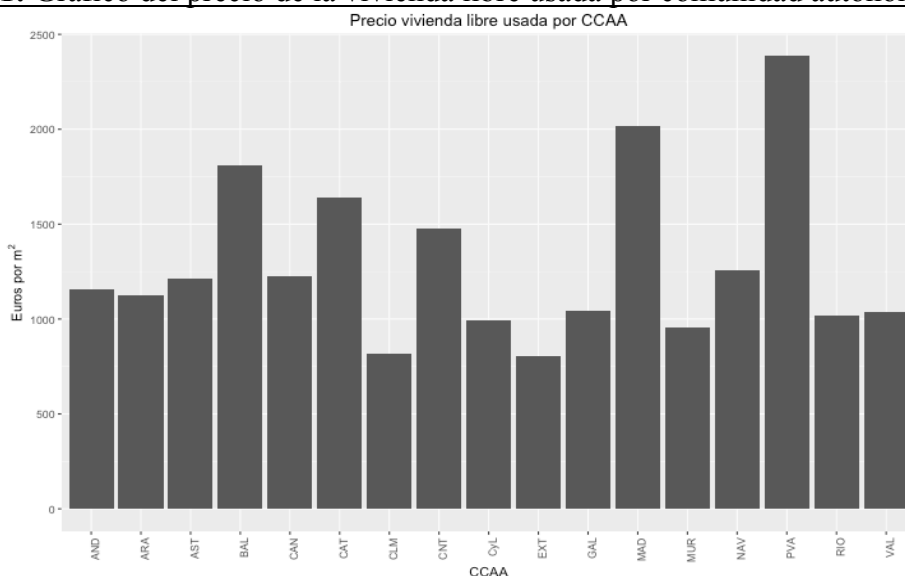
**Figura 3.10:** Gráfico del precio de la vivienda libre nueva por comunidad autónoma en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

- Precio vivienda libre usada: la comunidad autónoma con el precio de vivienda libre usada más elevado es el País Vasco, con 2385.2 euros el metro cuadrado. Por el contrario, la comunidad autónoma con el precio más bajo es Extremadura con 806.5 euros el metro cuadrado.

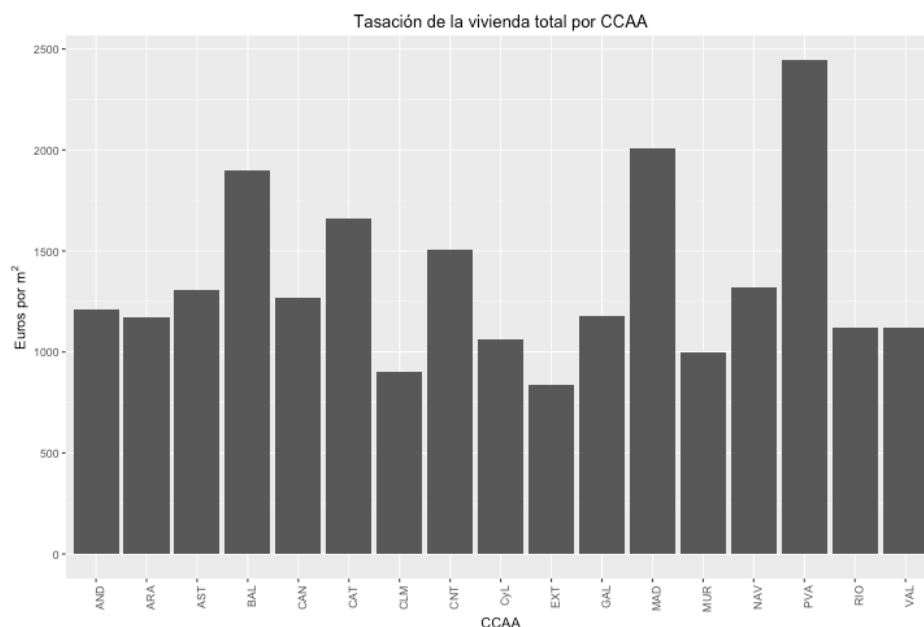
**Figura 3.11:** Gráfico del precio de la vivienda libre usada por comunidad autónoma en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

- Tasación vivienda total: la comunidad autónoma con la tasación de vivienda más elevada es, como ocurría con los precios, el País Vasco, con 2441.2 euros el metro cuadrado. La comunidad autónoma donde la vivienda está tasada a un menor precio es Extremadura, con 861.1 euros el metro cuadrado.

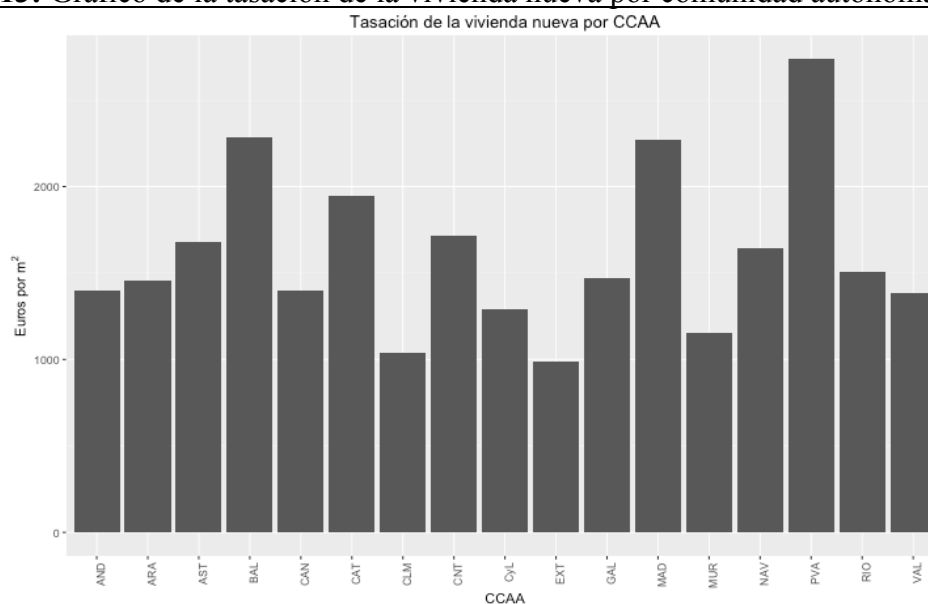
**Figura 3.12:** Gráfico de la tasación de la vivienda por comunidad autónoma en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

- Tasación vivienda nueva: la tasación de vivienda nueva más elevada la encontramos en el País Vasco, con una tasación de 2716 euros el metro cuadrado. La tasación de vivienda nueva de menos valor la encontramos en Extremadura, con 985.3 euros el metro cuadrado.

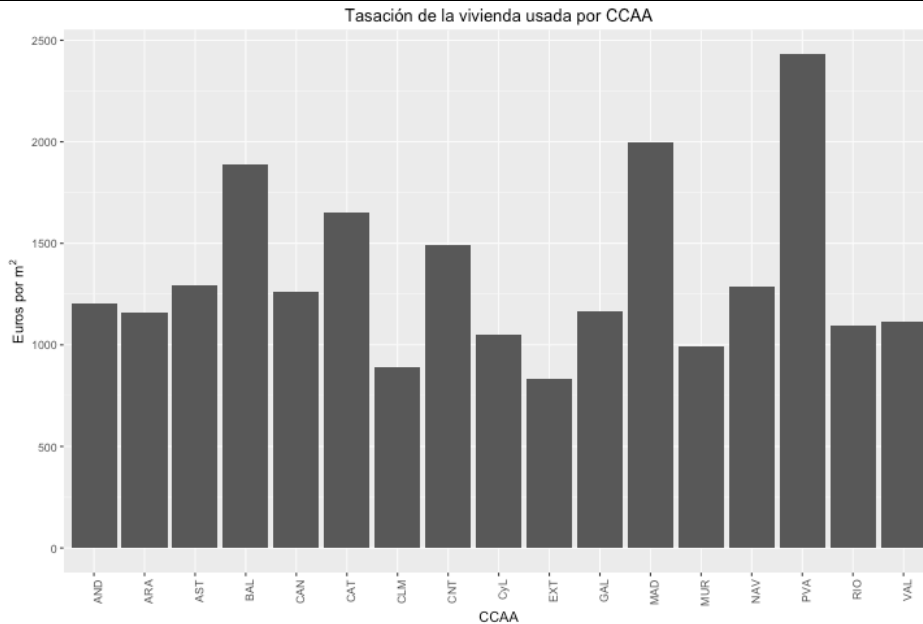
**Figura 3.13:** Gráfico de la tasación de la vivienda nueva por comunidad autónoma en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

- Tasación vivienda usada: la tasación de vivienda usada más elevada se encuentra en el País Vasco, con una tasación de 2425.6 euros el metro cuadrado. La tasación de vivienda nueva de menos valor se encuentra en Extremadura, con 855.9 euros el metro cuadrado.

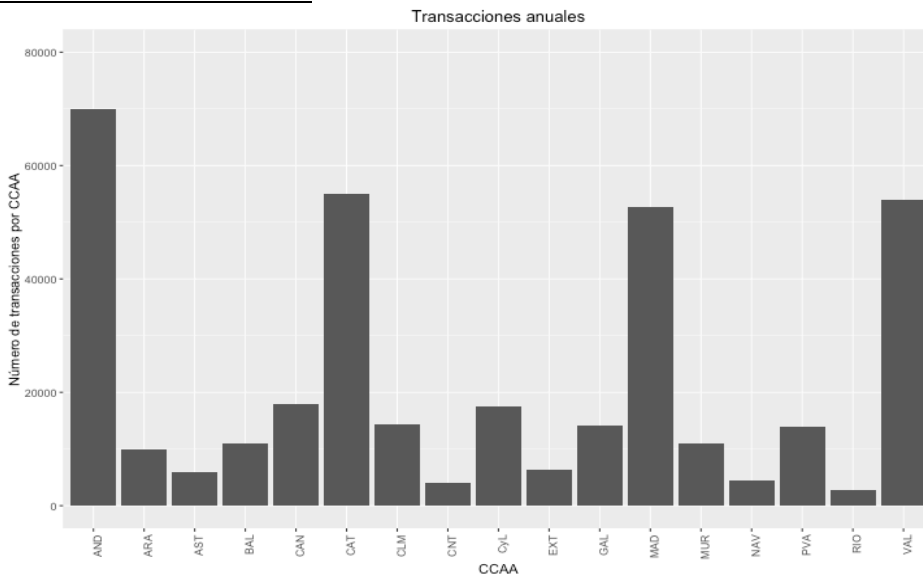
**Figura 3.14:** Gráfico de la tasación de la vivienda usada por comunidad autónoma en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

- Transacciones anuales: Encontramos cuatro comunidades autónomas con muchas más transacciones en el mercado de la vivienda que las otras. Estas son Andalucía, Cataluña, la Comunidad de Madrid y la Comunidad Valenciana, con 69 999, 54 948, 53 989 y 52 697 transacciones respectivamente. La media de la variable se encuentra en 21 453 transacciones, hecho que marca la diferencia entre comunidades.

**Figura 3.15:** Gráfico de las transacciones anuales en el mercado de la vivienda por comunidad autónoma en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

Dada la diferencia de población por comunidad autónoma, se puede observar el número de transacciones por habitante dentro de cada comunidad para poder comparar las comunidades.

**Tabla 3.11:** Descriptivos de la variable “Transacciones anuales” per cápita de los datos según comunidad autónoma.

	<i>Mín.</i>	<i>Q1</i>	<i>Q2</i>	<i>Media</i>	<i>Q3</i>	<i>Máx.</i>
<i>Transacciones anuales per cápita</i>	0.005107	0.006795	0.007308	0.007473	0.008331	0.010787

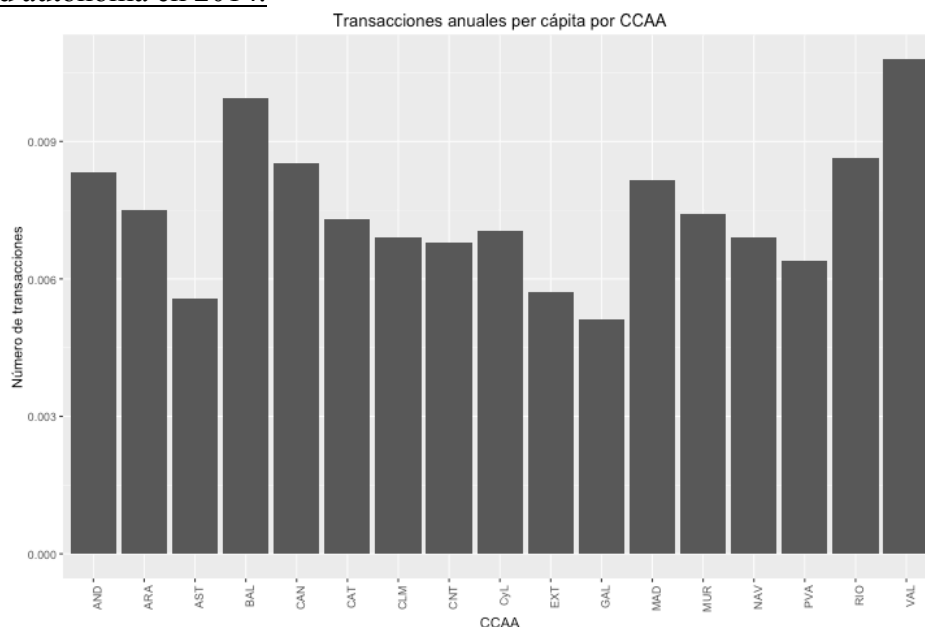
Fuente: Ministerio de Fomento. Elaboración propia.

- Transacciones anuales por habitante: gráficamente (ver Figura 3.16), se ve como en el número de transacciones cuando se pondera por población no genera las grandes diferencias que genera el número de transacciones por si sola.

En este caso, la comunidad autónoma con más transacciones por habitante es la Comunidad Valenciana, con 0.0107 transacciones por habitante y la que menos Galicia, con 0.0051 transacciones por habitante.

Las cuatro comunidades que en la variable anterior tomaban el liderazgo de las transacciones, en este caso las encontramos en la posición segunda (Baleares), quinta (Andalucía), sexta (Comunidad de Madrid) y novena (Cataluña).

**Figura 3.16:** Gráfico de las transacciones anuales per cápita en el mercado de la vivienda por comunidad autónoma en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

Otro dato importante a tener en cuenta es la diferencia entre el precio de mercado y el valor de tasación de la vivienda. Con los estadísticos descriptivos presentados a

continuación, podemos comprobar que, actualmente, los valores de tasación de la vivienda son más elevados que los precios de mercado.

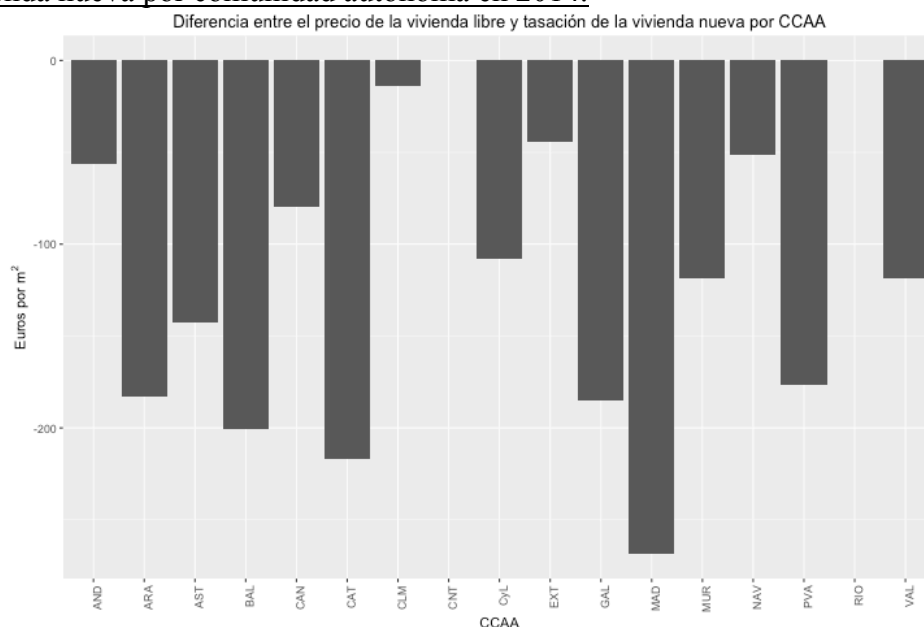
**Tabla 3.12:** Descriptivos de diferencia entre las variables precio de la vivienda y tasación para la vivienda total, nueva y usada de los datos según comunidad autónoma.

	<i>Mín.</i>	<i>Q1</i>	<i>Q2</i>	<i>Media</i>	<i>Q3</i>	<i>Máx.</i>
<i>Diferencia entre precio vivienda total y tasación de la vivienda total</i>	0	0	0	0	0	0
<i>Diferencia entre precio vivienda nueva y tasación de la vivienda nueva</i>	-398.8	-192.6	-142.8	-150.4	-80.0	-13.8
<i>Diferencia entre precio vivienda usada y tasación de la vivienda usada</i>	-118.70	-76.50	-44.80	-49.25	-33.10	16.10

Fuente: Ministerio de Fomento. Elaboración propia.

- Diferencia entre precio vivienda libre y tasación vivienda total: según los datos del Atlas Digital, la diferencia entre el precio de la vivienda nueva y la tasación de esta en 2014 era del mismo valor.
- Diferencia entre precio vivienda libre y tasación vivienda nueva: Castilla – La Mancha es la comunidad autónoma con menor diferencia entre el valor de tasación de la vivienda nueva y el precio de esta, siendo el valor de tasación 13.8 euros el metro cuadrado más elevado que el precio. La comunidad autónoma donde esta diferencia es mayor es la Comunidad de Madrid, con una diferencia de 268.6 euros el metro cuadrado.

**Figura 3.17:** Gráfico de la diferencia entre el precio de la vivienda libre nueva y la tasación de la vivienda nueva por comunidad autónoma en 2014.

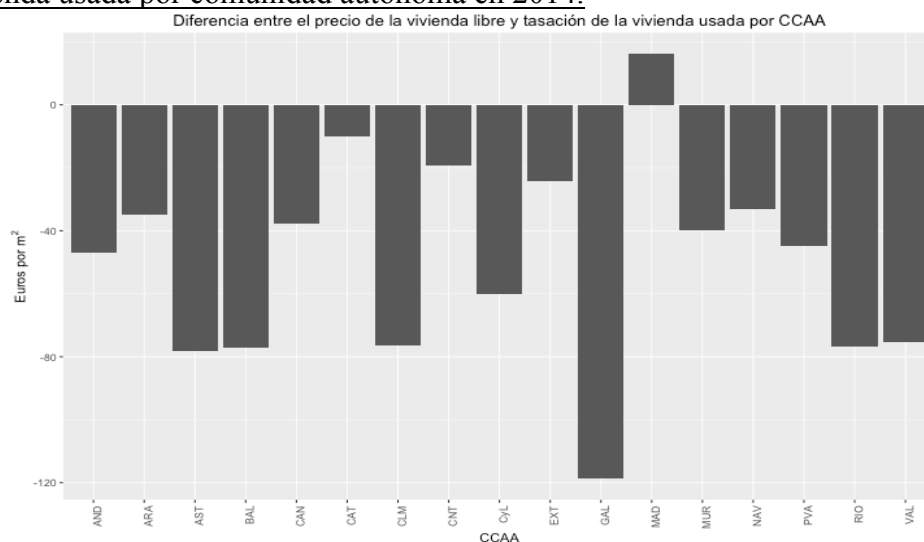


Fuente: Ministerio de Fomento. Elaboración propia.

- Diferencia entre precio vivienda libre y tasación vivienda usada: en este caso, la Comunidad de Madrid tiene el precio de la vivienda usada superior al valor de

tasación, siendo la diferencia de 16.10 euros el metro cuadrado. Las demás comunidades autónomas tienen un valor de tasación superior al precio, siendo el máximo de 118.7 euros el metro cuadrado en Galicia.

**Figura 3.18:** Gráfico de la diferencia entre el precio de la vivienda libre usada y la tasación de la vivienda usada por comunidad autónoma en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

### 3.3.2.3. Comparativa del año 2014 con 2004

Una descriptiva de los datos del año 2004 lo podemos encontrar a continuación.

**Tabla 3.13:** Descriptivos de las variables referentes al precio de la vivienda y a su tasación en 2004 de los datos según comunidad autónoma.

	<i>Mín.</i>	<i>Q1</i>	<i>Q2</i>	<i>Media</i>	<i>Q3</i>	<i>Máx.</i>
<i>Precio vivienda total</i>	739.3	1150.2	1301.6	1396.7	1439.9	2387.1
<i>Precio vivienda nueva</i>	821.1	1232.3	1420.5	1474.0	1523.9	2293.1
<i>Precio vivienda usada</i>	706.3	1094.0	1206.9	1351.1	1399.5	2410.2
<i>Tasación vivienda total</i>	739.3	1150.2	1301.6	1396.7	1439.9	2387.1
<i>Tasación vivienda nueva</i>	821.1	1232.3	1420.5	1474.0	1523.9	2293.1
<i>Tasación vivienda usada</i>	706.3	1094.0	1206.9	1351.1	1399.5	2410.2

Fuente: Ministerio de Fomento. Elaboración propia.

Pero lo que realmente interesa es saber la diferencia entre 2004 y el año 2014, para ver como ha evolucionado el mercado de la vivienda la última década.

**Tabla 3.14:** Descriptivos de la diferencia de las variables referentes al precio de la vivienda ya la tasación entre 2004 y 2014 de los datos según comunidad autónoma.

	<i>Mín.</i>	<i>Q1</i>	<i>Q2</i>	<i>Media</i>	<i>Q3</i>	<i>Máx.</i>
<i>Precio vivienda total</i>	-379.00	-113.90	-80.50	-43.49	64.80	221.30
<i>Precio vivienda nueva</i>	-291.40	-122.60	-47.20	-14.76	120.00	295.80
<i>Precio vivienda usada</i>	-396.30	-134.90	-101.50	-58.21	75.10	189.10
<i>Tasación vivienda total</i>	-379.00	-113.90	-80.50	-43.49	64.80	221.30
<i>Tasación vivienda nueva</i>	821.1	1232.3	1420.5	1474.0	1523.9	2293.1
<i>Tasación vivienda usada</i>	706.3	1094.0	1206.9	1351.1	1399.5	2410.2

Fuente: Ministerio de Fomento. Elaboración propia.

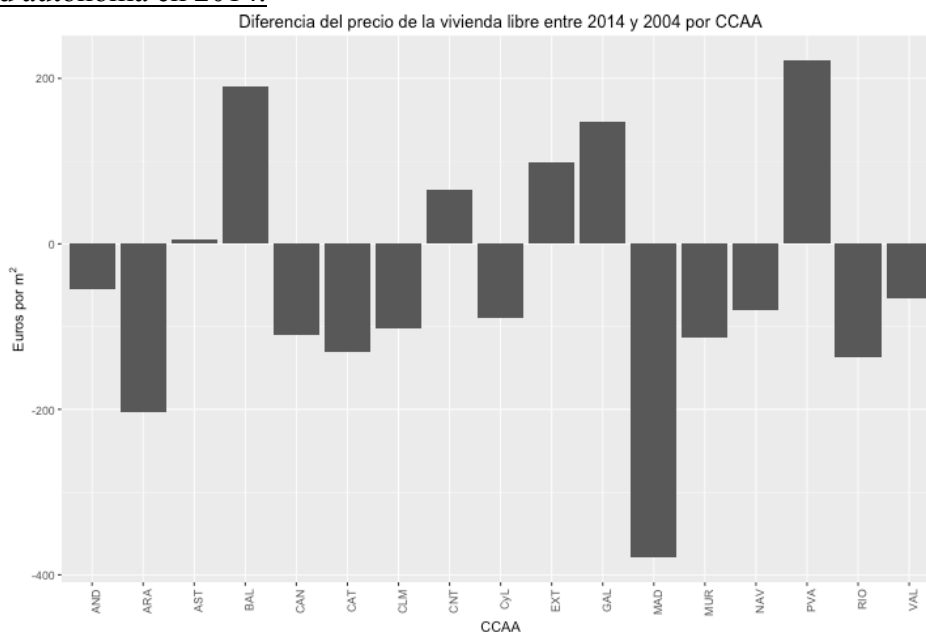
Destacan las comunidades autónomas de Baleares, Cantabria, Extremadura, Galicia y el País Vasco, donde los precios de la vivienda libre y el valor de la tasación, en el año 2014, estaban por encima del valor que tenían en 2004.

En el caso de Castilla - La Mancha y Asturias, el valor de tasación en 2014 era superior al de 2004, mientras que los precios de vivienda eran más bajos en 2014 que en 2004.

En las demás comunidades, el precio de la vivienda libre y el valor de tasación de la vivienda eran más elevados en 2004 respecto a 2014.

- Diferencia entre el precio de la vivienda libre total entre 2014 y 2004:

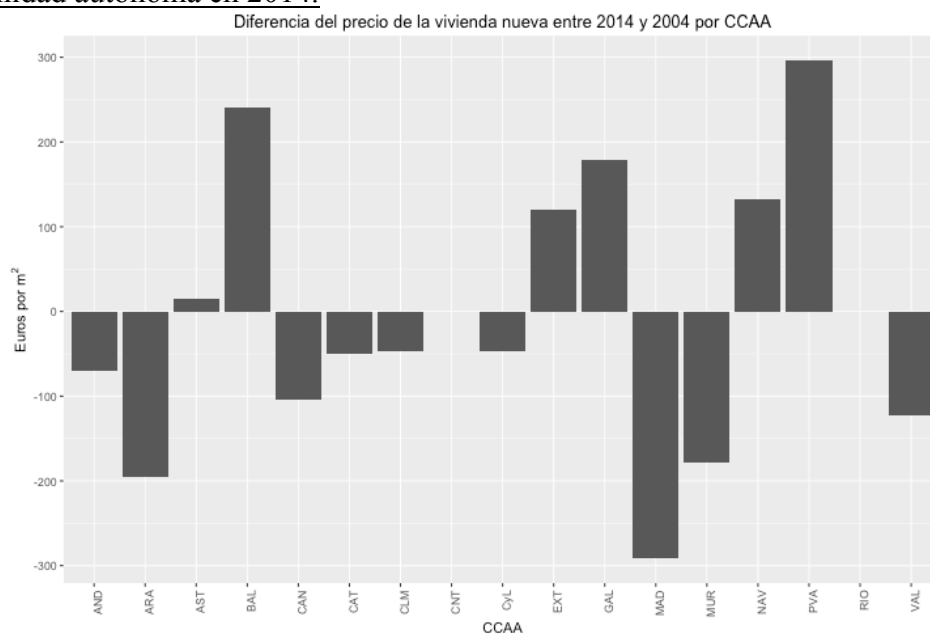
**Figura 3.19:** Gráfico de la diferencia del precio de la vivienda libre entre 2014 y 2004 por comunidad autónoma en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

- Diferencia entre el precio de la vivienda libre nueva entre 2014 y 2004:

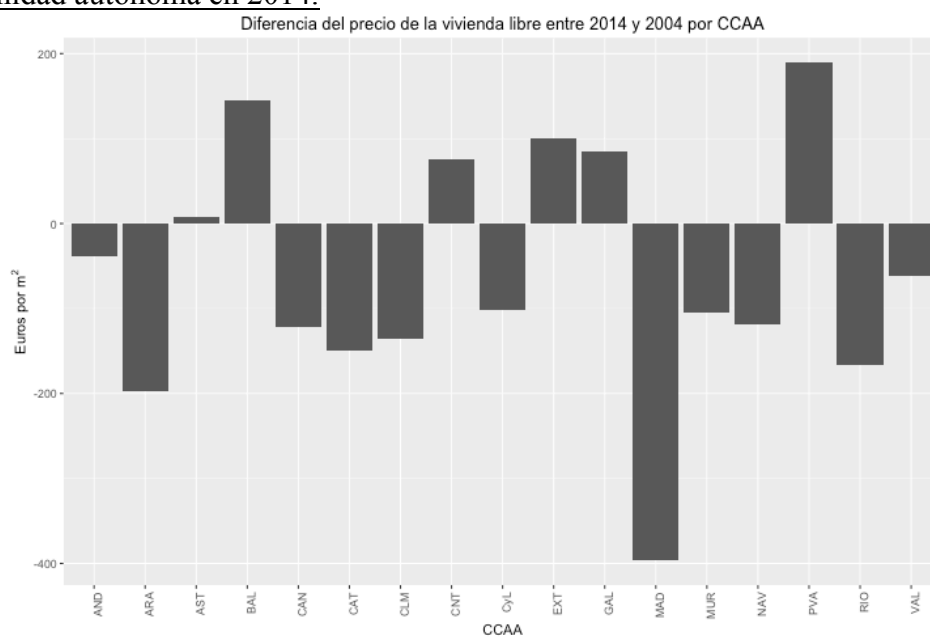
**Figura 3.20:** Gráfico de la diferencia del precio de la vivienda libre nueva entre 2014 y 2004 por comunidad autónoma en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

- Diferencia entre el precio de la vivienda libre usada entre 2014 y 2004:

**Figura 3.21:** Gráfico de la diferencia del precio de la vivienda libre usada entre 2014 y 2004 por comunidad autónoma en 2014.

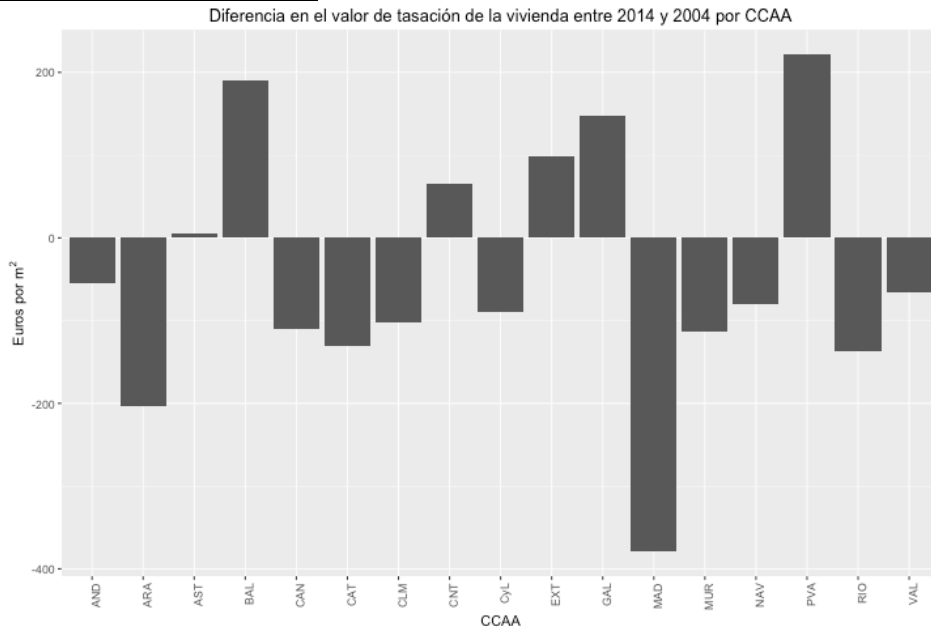


Fuente: Ministerio de Fomento. Elaboración propia.



- Diferencia entre el valor de tasación de vivienda total entre 2014 y 2004:

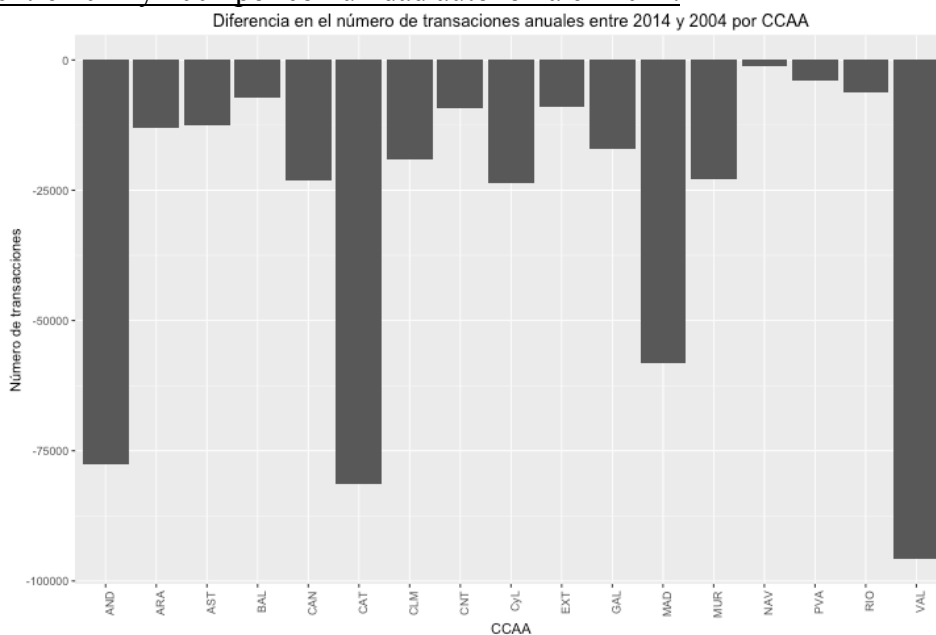
**Figura 3.22:** Gráfico de la diferencia de la tasación de la vivienda libre entre 2014 y 2004 por comunidad autónoma en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

- Diferencia entre el número de transacciones anuales entre 2014 y 2004: el número de transacciones en el mercado de vivienda ha decrecido en el año 2014 respecto al año 2004.

**Figura 3.23:** Gráfico de la diferencia del número de transacciones en el mercado de la vivienda entre 2014 y 2004 por comunidad autónoma en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

### 3.3.2.4. Descriptiva de las variables del parque de vivienda

Se presentan los principales estadísticos descriptivos (mínimo, primer cuartil, mediana, media, tercer cuartil y máximo) para las cuatro variables numéricas del parque de la vivienda que hacen referencia a la densidad de viviendas por kilómetro cuadrado, que son “densidad de viviendas familiares”, “densidad de viviendas familiares principales”, “densidad de viviendas secundarias” y “densidad de viviendas vacías”, con datos del censo de 2011.

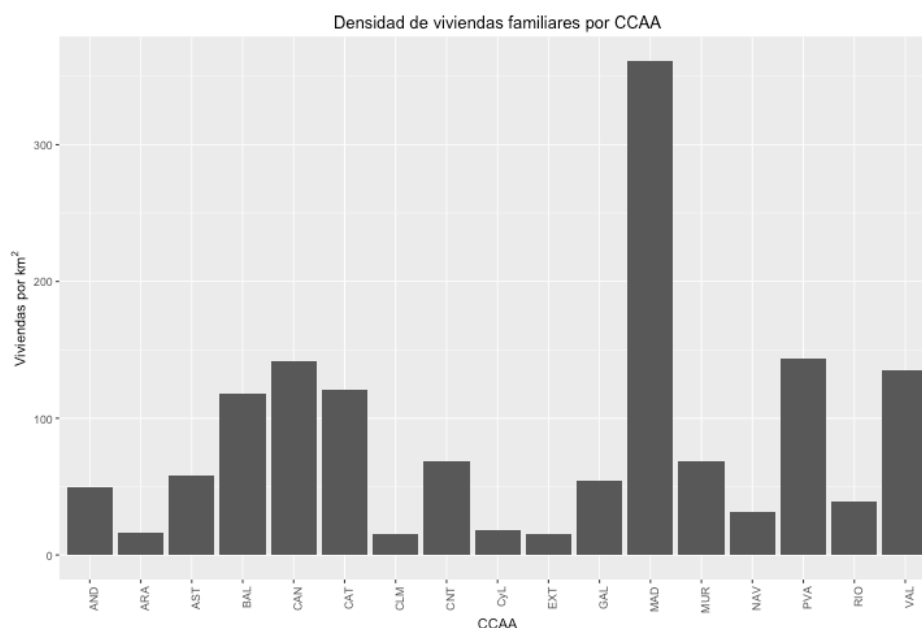
**Tabla 3.15:** Descriptivos referentes a la densidad de viviendas según tipo de los datos según comunidad autónoma.

	<i>Mín.</i>	<i>Q1</i>	<i>Q2</i>	<i>Media</i>	<i>Q3</i>	<i>Máx.</i>
<i>Densidad viviendas familiares</i>	15.60	31.50	57.90	85.58	120.50	360.90
<i>Densidad viviendas familiares principales</i>	9.90	25.40	43.20	64.86	86.10	307.90
<i>Densidad viviendas secundarias</i>	2.50	4.50	7.20	10.18	14.70	28.20
<i>Densidad viviendas vacías</i>	2.10	3.60	9.00	10.55	14.00	32.80

Fuente: Ministerio de Fomento. Elaboración propia.

- Densidad de viviendas familiares: la Comunidad de Madrid es la comunidad autónoma con mayor densidad de viviendas familiares, con 360.9 viviendas por kilómetro cuadrado. La siguen las comunidades autónomas de costa e islas, mientras que las de interior son las que menor densidad de viviendas familiares tienen.

**Figura 3.24:** Gráfico de la densidad de viviendas familiares por comunidad autónoma en 2014.

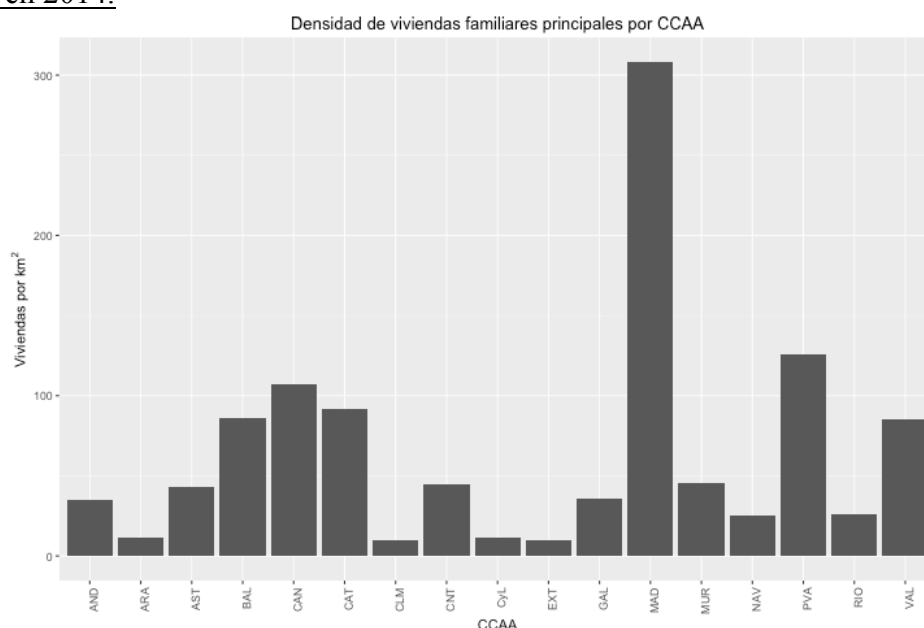


Fuente: Ministerio de Fomento. Elaboración propia.

- Densidad de viviendas familiares principales: cuando hablamos de viviendas familiares principales, la distribución gráfica de la variable podemos observar que

sigue un comportamiento parecido al de la anterior variable, pero siendo en este caso los valores menores que en la variable anterior.

**Figura 3.25:** Gráfico de la densidad de viviendas familiares principales por comunidad autónoma en 2014.



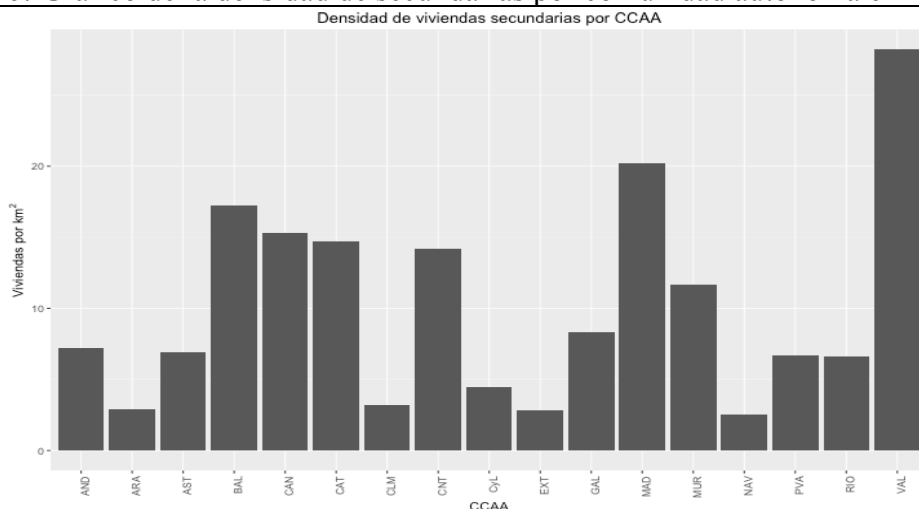
Fuente: Ministerio de Fomento. Elaboración propia.

- Densidad de viviendas secundarias: cuando nos centramos en las viviendas secundarias, vemos como la Comunidad Valenciana es la que más viviendas secundarias por kilómetro cuadrado tiene, con 28.2 viviendas secundarias por kilómetro cuadrado.

Le siguen la Comunidad de Madrid y regiones de costa e islas como Baleares, Canarias, Cataluña, Cantabria y Murcia.

La comunidad con menos densidad de viviendas secundarias es Navarra con 2.5 viviendas secundarias por kilómetro cuadrado.

**Figura 3.26:** Gráfico de la densidad de secundarias por comunidad autónoma en 2014.

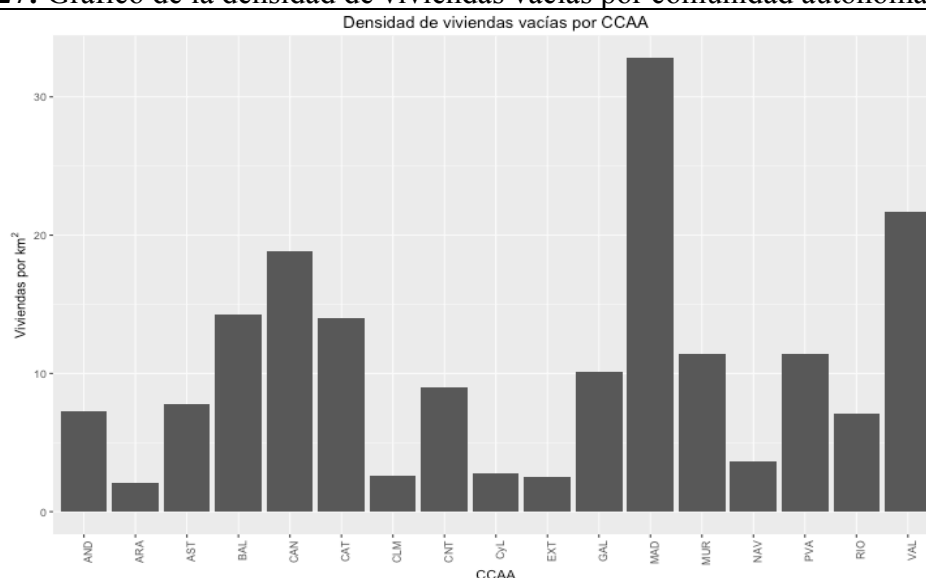


Fuente: Ministerio de Fomento. Elaboración propia.

- Densidad de viviendas vacías: la Comunidad de Madrid es la comunidad que presenta una mayor densidad de viviendas vacías, con 32.8 viviendas vacías por kilómetro cuadrado.

Destaca que las comunidades con menos densidad de viviendas vacías son las comunidades autónomas de interior.

**Figura 3.27:** Gráfico de la densidad de viviendas vacías por comunidad autónoma en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

Se presentan los principales estadísticos descriptivos (mínimo, primer cuartil, mediana, media, tercer cuartil y máximo) para las tres variables numéricas del parque de la vivienda que hacen referencia al número de viviendas, que son el número de viviendas familiares principales, número de viviendas secundarias y número de viviendas vacías, con datos del censo de 2011.

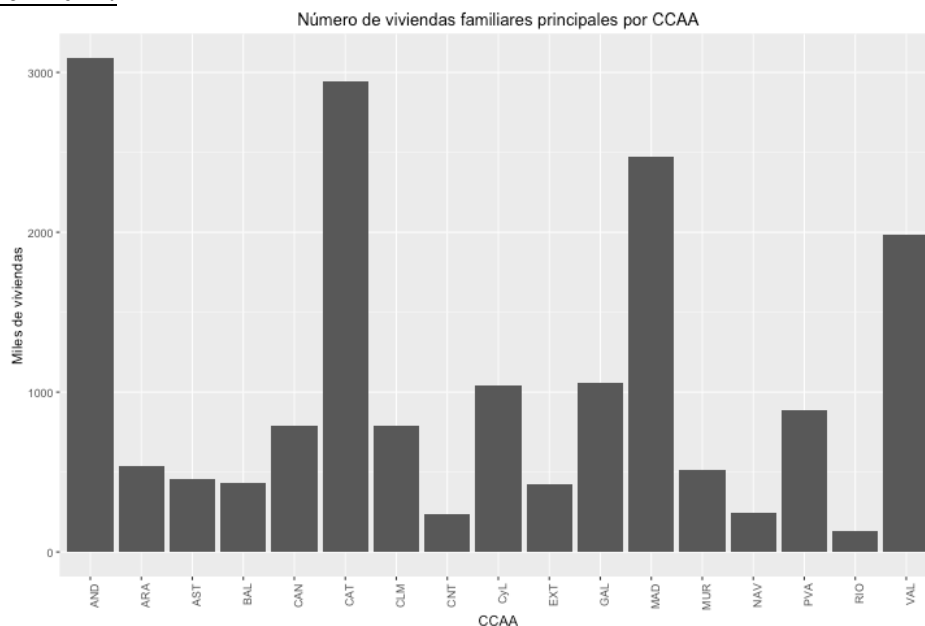
**Tabla 3.16:** Descriptivos referentes número de viviendas según tipo de los datos según comunidad autónoma.

	<i>Mín.</i>	<i>Q1</i>	<i>Q2</i>	<i>Media</i>	<i>Q3</i>	<i>Máx.</i>
<i>Viviendas familiares principales</i>	129984	429737	787516	1060831	1059233	3087222
<i>Viviendas secundarias</i>	24540	74697	132216	216527	254768	655137
<i>Viviendas vacías</i>	35466	81003	129117	202392	263279	637221

Fuente: Ministerio de Fomento. Elaboración propia.

- Número de viviendas familiares principales: las comunidades autónomas con más población tienen mayor número de viviendas familiares principales.

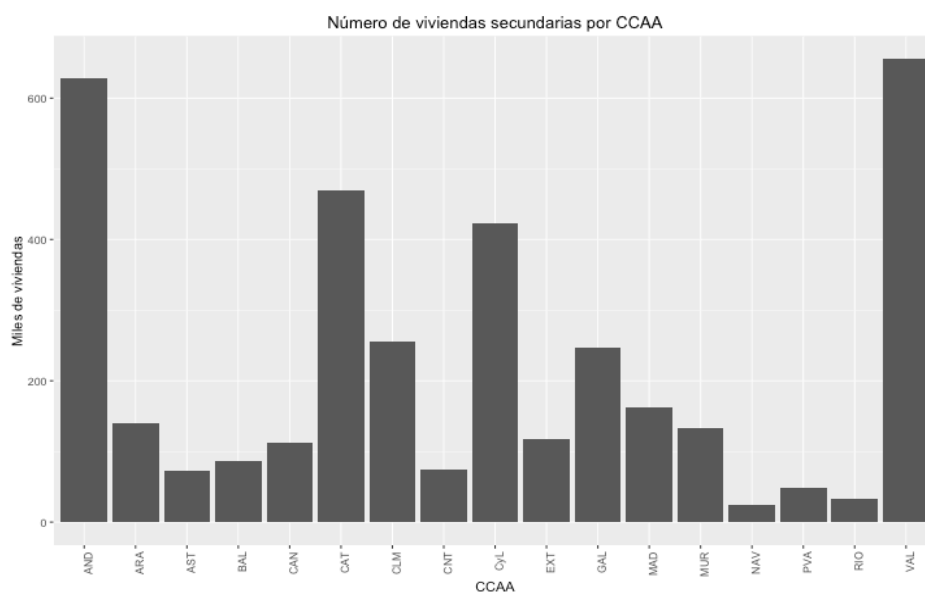
**Figura 3.28:** Gráfico del número de viviendas familiares principales por comunidad autónoma en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

- Número de viviendas secundarias: las comunidades con más viviendas secundarias son la Comunidad Valenciana y Andalucía, con 655 137 y 628 703 viviendas secundarias respectivamente. La comunidad con menos viviendas secundarias es Navarra, con 24 540 viviendas secundarias.

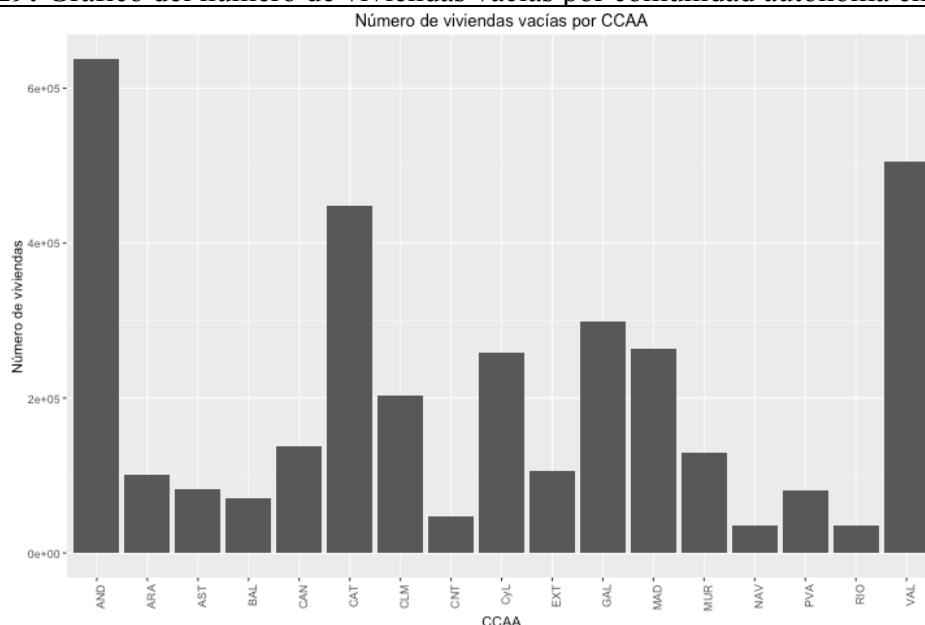
**Figura 3.28:** Gráfico del número de viviendas secundarias por comunidad autónoma en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

- Número de viviendas vacías: las comunidades con más viviendas vacías son Andalucía y la Comunidad Valenciana, con 637 221 y 505 029 viviendas vacías respectivamente, mientras que las comunidades autónomas con menos viviendas vacías son Navarra y La Rioja, con 35 466 y 35 745 viviendas vacías respectivamente.

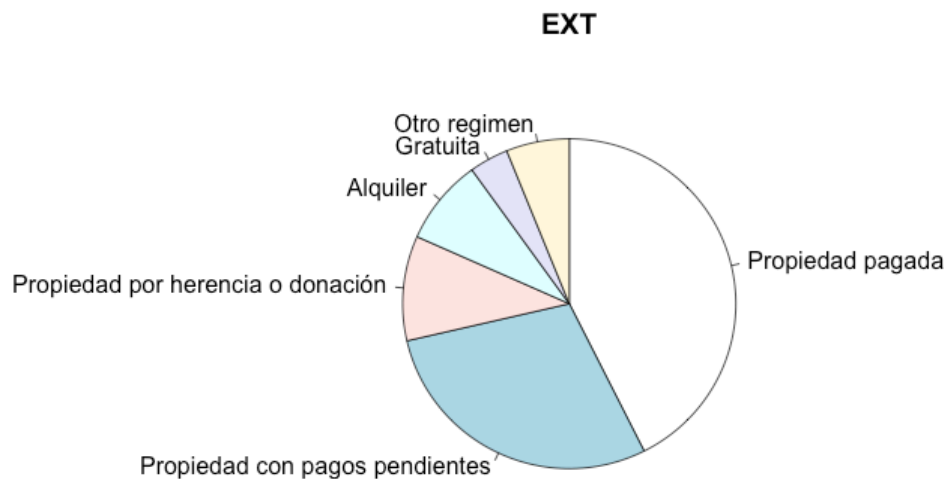
**Figura 3.29:** Gráfico del número de viviendas vacías por comunidad autónoma en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

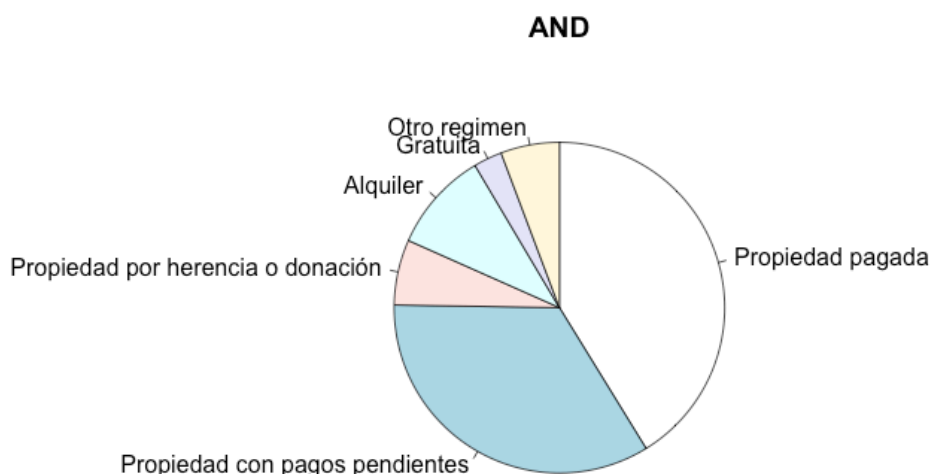
- Régimen de tenencia: las distintas comunidades autónomas tienen diferentes distribuciones en los regímenes de tenencia, aunque todas parecidas, con una mayor presencia de los regímenes de tenencia relacionados con la propiedad (superiores al 75%). Para ello, se presentan gráficamente tres ejemplos, una región de costa, otra de interior y, finalmente, la región de la capital del Estado.

**Figura 3.30:** Gráfico de los regímenes de tenencia de las viviendas de Extremadura en 2014.



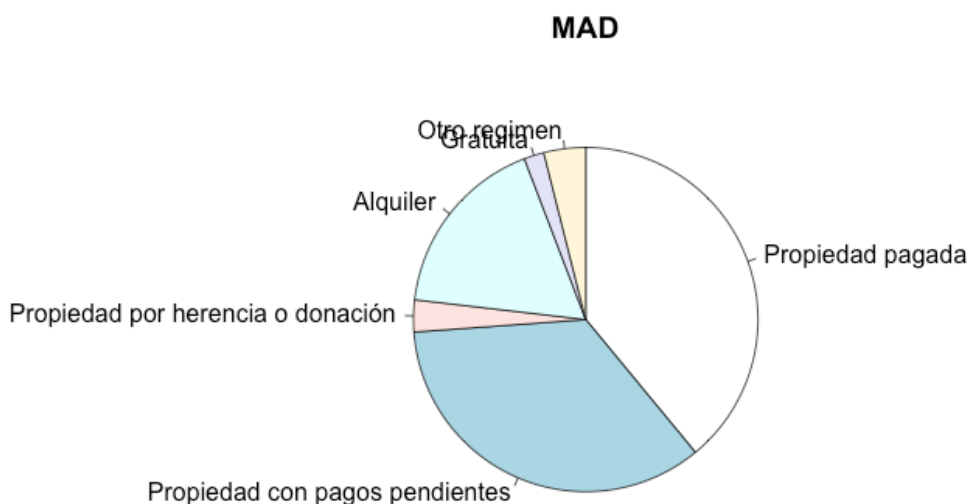
Fuente: Ministerio de Fomento. Elaboración propia.

**Figura 3.31:** Gráfico de los regímenes de tenencia de las viviendas de Andalucía en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

**Figura 3.32:** Gráfico de los regímenes de tenencia de las viviendas en la Comunidad de Madrid en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

### 3.3.2.5. Descriptiva de las variables de la ocupación del suelo

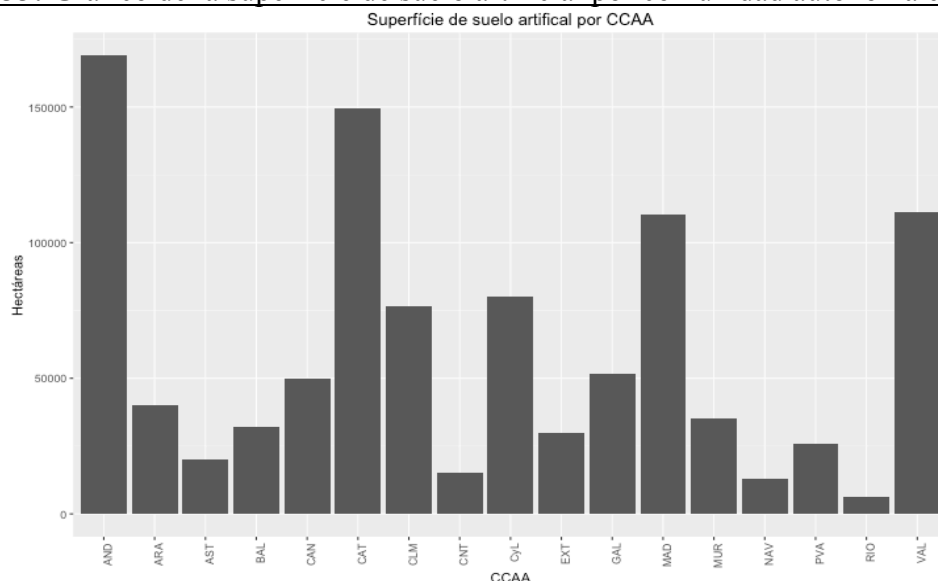
La variable “superficie de suelo artificial” delimita la cantidad de terreno que puede ser destinado a la construcción con distintas finalidades, entre ellas la vivienda. Los valores de esta variable demuestran que el tamaño de cada región no necesariamente está relacionado con la superficie de suelo artificial. A continuación, se presentan los descriptivos básicos de esta variable en el año 2006.

**Tabla 3.16:** Descriptivos de la variable “superficie de suelo artificial” de los datos según comunidad autónoma.

	<i>Mín.</i>	<i>Q1</i>	<i>Q2</i>	<i>Media</i>	<i>Q3</i>	<i>Máx.</i>
<i>Superficie de suelo artificial</i>	6156	25935	40260	59756	80294	168922

Fuente: Ministerio de Fomento. Elaboración propia.

**Figura 3.33:** Gráfico de la superficie de suelo artificial por comunidad autónoma en 2014.

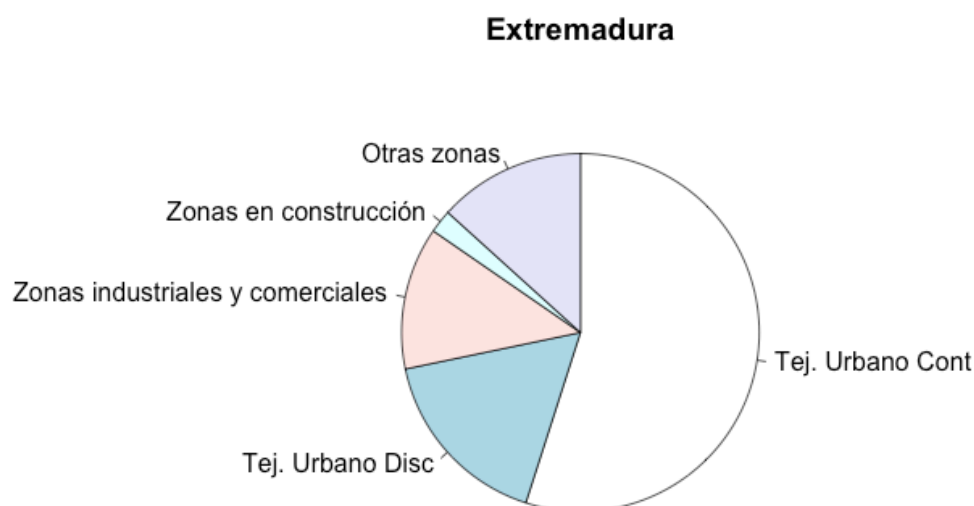


Fuente: Ministerio de Fomento. Elaboración propia.

Se puede observar como las comunidades con mayor superficie artificial son Andalucía y Cataluña, siendo estas las comunidades más pobladas.

La distribución del suelo artificial con los datos del año 2006 estaba repartida, básicamente, entre el *tejido urbano continuo*, *tejido urbano discontinuo*, *zonas industriales y comerciales*, *zonas en construcción* y el conjunto de las demás categorías. Para ver esta distribución, se presenta en gráficos de sectores las distribuciones del suelo en las tres comunidades donde anteriormente se ha presentado los regímenes de tenencia.

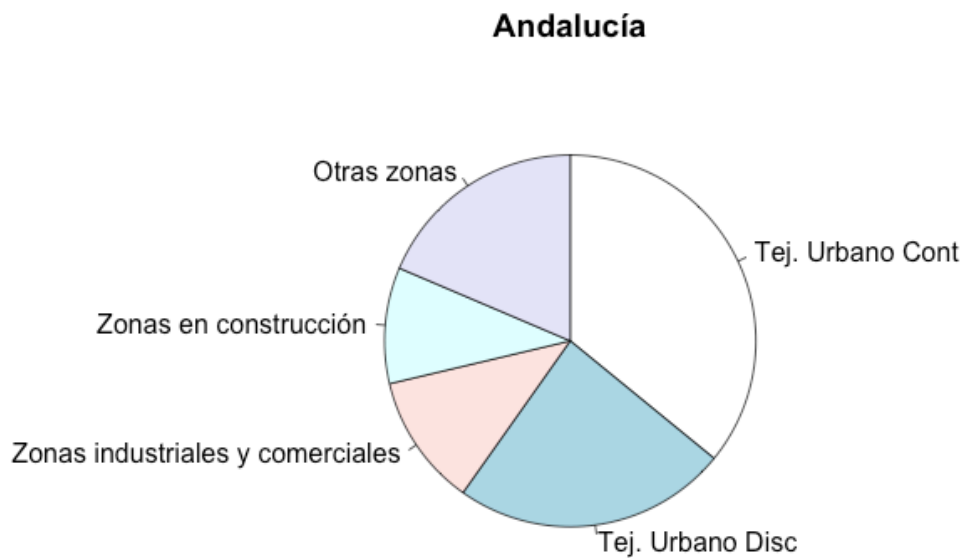
**Figura 3.34:** Gráfico de la distribución del suelo artificial en Extremadura en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.



**Figura 3.35:** Gráfico de la distribución del suelo artificial en Andalucía en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

**Figura 3.36:** Gráfico de la distribución del suelo artificial en la Comunidad de Madrid en 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

### 3.4. Tratamiento de datos *missing* y interpolación y extrapolación de datos

Los datos obtenidos del Atlas Digital son los siguientes:

- Datos demográficos para el período 2004 a 2014.
- Datos del mercado de vivienda para el período 2004 a 2014.
- Datos del parque de la vivienda a partir de los censos de 2001 y 2011.
- Datos de la ocupación del suelo para los años 2000 y 2006.

El primer tratamiento aplicado sobre los datos es la imputación de valores a los valores *missing*. Esta imputación se hace a través de la mediana de la variable en aquel año, dado

que la mediana es un estimador robusto, es decir, no queda afectada por los valores extremos de las variables ni por cambios en los datos.

Dada la falta de datos para algunos años en las variables del parque de la vivienda y de la ocupación del suelo, se decide aplicar el método de la interpolación geométrica para obtener los valores de las variables para los años entre los cuales se tienen datos. Para eso, calculamos la razón a través del valor de la variable en los dos momentos temporales de los cuales disponemos de datos.

$$r = \left( \frac{a_2}{a_1} \right)$$

donde  $a_2$  es el valor de la variable en el momento temporal final y  $a_1$  es el valor de la variable en el momento temporal inicial.

Con la razón, se calculan cada uno de los valores de la variable para la cual no se tienen datos con la fórmula:

$$a_n = a_1 r^{\frac{n}{T}}$$

donde  $n$  es el momento temporal entre el momento inicial y el momento final y  $T$  es el período temporal entre el momento inicial y final.

Para los datos de fuera del intervalo, usamos el método de la extrapolación a través de la misma fórmula de la interpolación, siendo  $n$  el número de períodos temporales entre el momento temporal para el cual se quieren imputar los valores y el momento inicial.

Estos métodos se aplican para la obtención de los datos para los años 2004 y 2005 y 2007 a 2014 para las variables de la ocupación del suelo y para los años 2004 a 2010 y 2012 a 2014 para las variables del parque de la vivienda.

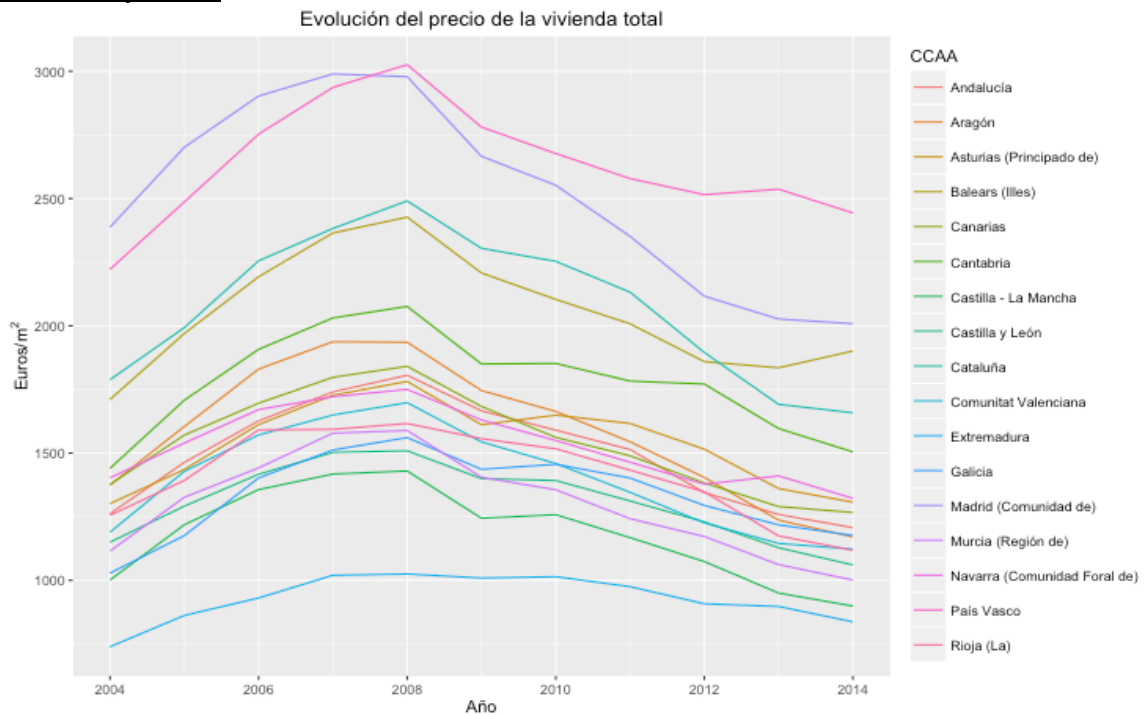
También, se decide transformar las variables *zona portuaria* y *aeropuertos* en binarias según si su valor es 0 o mayor a 0, para marcar aquellas regiones que son de costa (con la variable *zona portuaria*) y que tienen aeropuerto, como factor diferencial con otras regiones.

### 3.5. Descriptiva de los datos tratados

Con los datos tratados, se crea una base de datos con los valores anuales de cada variable por cada región y año. Así, se puede hacer un análisis exploratorio temporal.

La primera variable interesante a observar es el precio de la vivienda en venta, tanto usada como nueva. En esta variable se puede observar el comportamiento visto hasta ahora en otras variables, con crecimiento hasta el año 2007 o 2008 y a partir de este punto el precio decrece. Destaca la diferencia de precio entre algunas comunidades autónomas, hecho que se analizará más adelante.

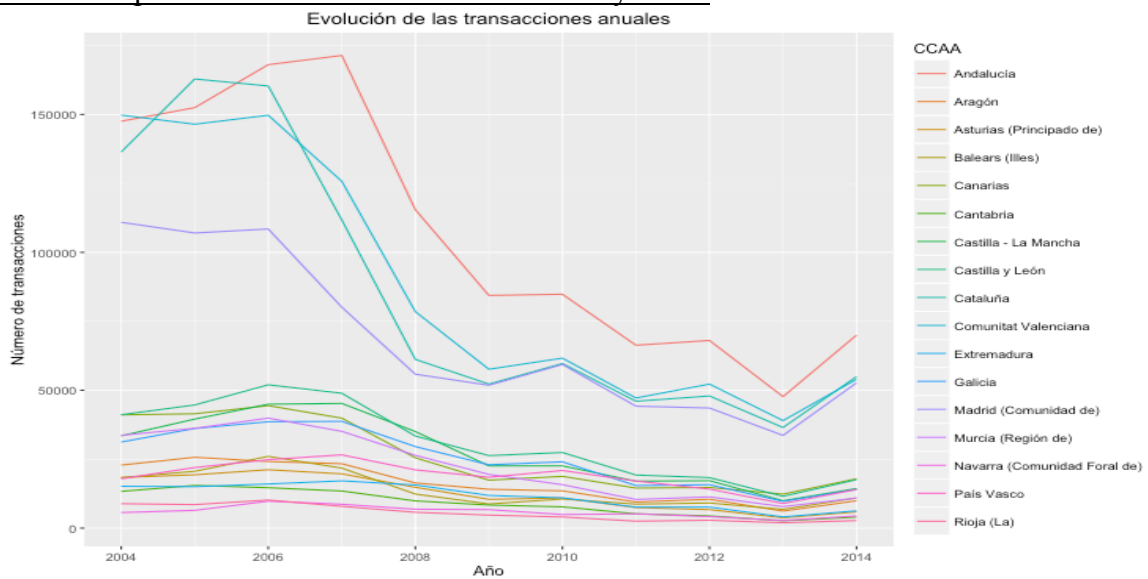
**Figura 3.37:** Gráfico de la evolución del precio de la vivienda por comunidad autónoma entre 2004 y 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

Siguiendo con la tendencia, se observa como el número de transacciones en el mercado de la vivienda encuentra su máximo, según cada comunidad autónoma, entre 2006 y 2008. A partir de este momento, el número de transacciones anuales decrece, sobretodo en cuatro regiones: Andalucía, Cataluña, Comunidad Valenciana y la Comunidad de Madrid. Destacar que estas regiones son de las que más población tienen, aunque la actividad de los habitantes de cada comunidad autónoma no es proporcional a su población.

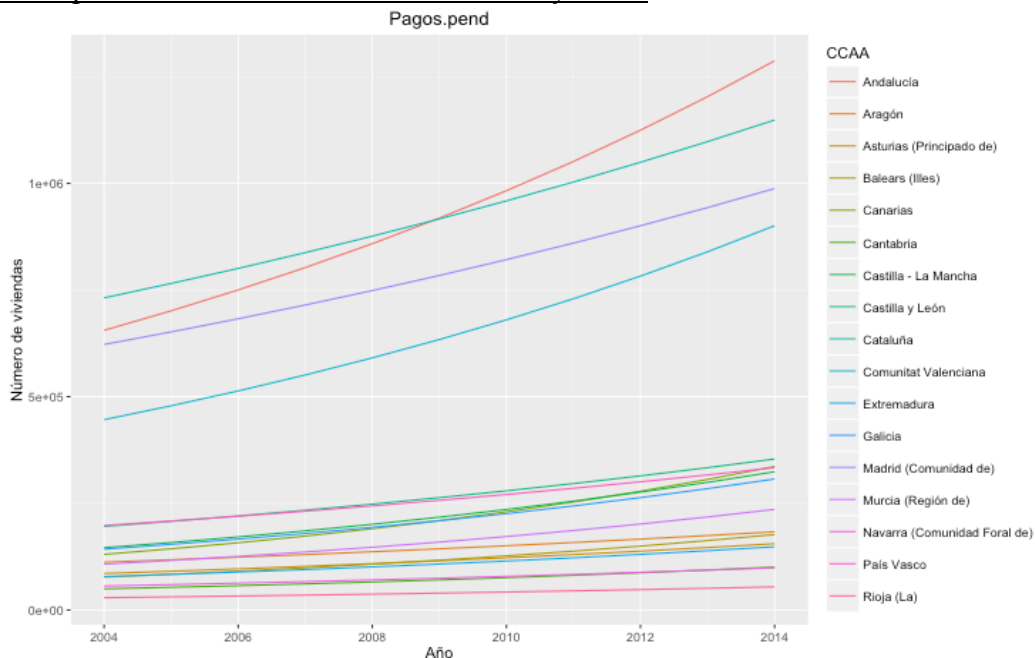
**Figura 3.38:** Gráfico de la evolución del número de transacciones anuales en el mercado de la vivienda por comunidad autónoma entre 2004 y 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

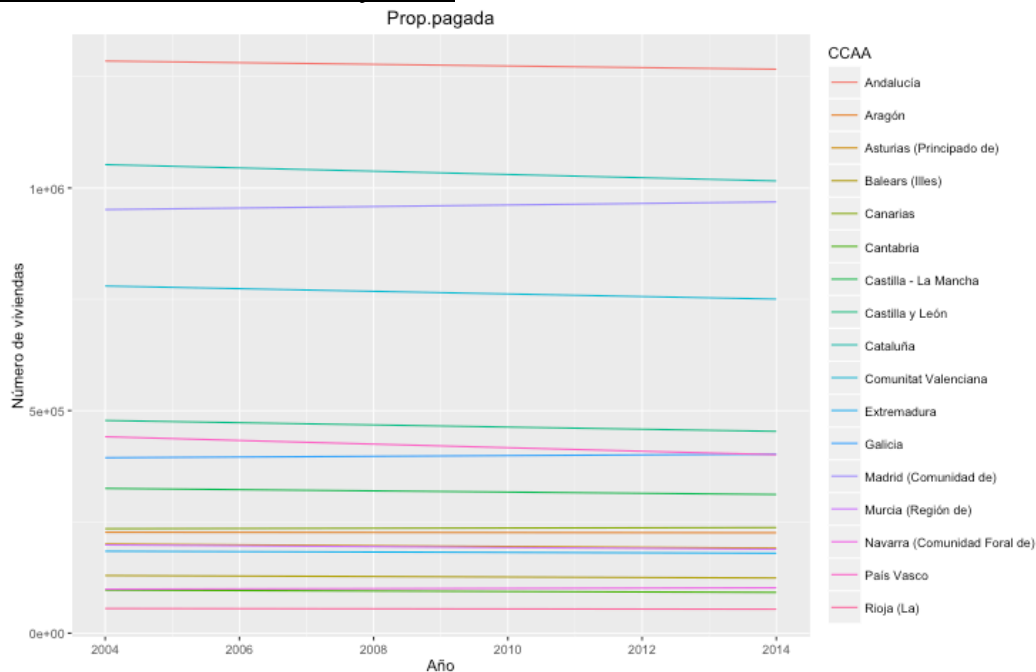
Por otra parte, se observa un importante crecimiento en el número de viviendas en propiedad sobre las que aún existe una deuda sobre ellas. El crecimiento es más importante en las comunidades en las que se observaba un mayor número de transacciones en los años previos a la crisis actual.

**Figura 3.39:** Gráfico de la evolución del número de viviendas en propiedad con pagos pendientes por comunidad autónoma entre 2004 y 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

**Figura 3.40:** Gráfico de la evolución del número de viviendas en propiedad pagada por comunidad autónoma entre 2004 y 2014.

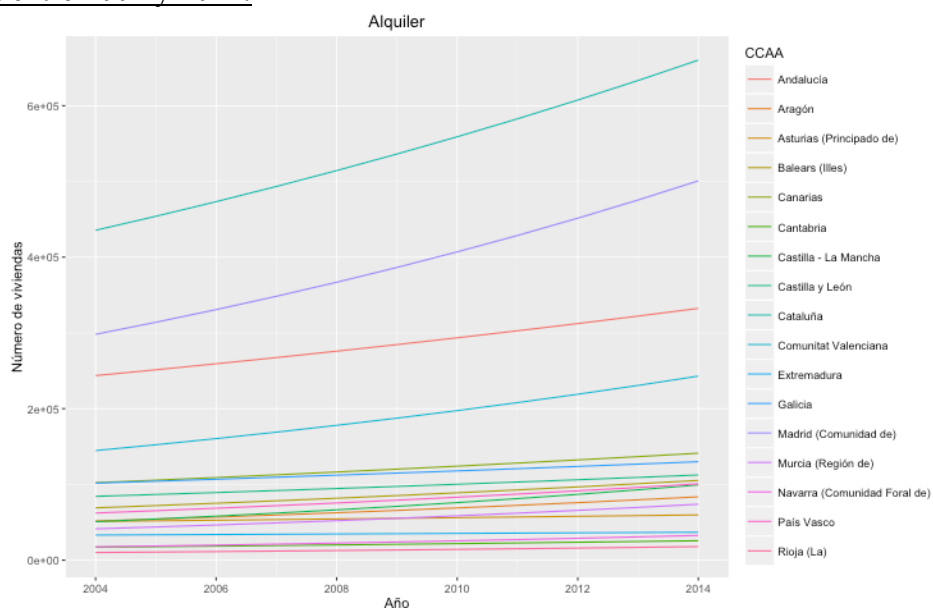


Fuente: Ministerio de Fomento. Elaboración propia.

Por su parte, las viviendas en propiedad sin deudas sobre ellas han mantenido su número estable a lo largo de la última década.

Además, y como se había visto en el análisis de la situación actual del mercado de la vivienda en España, el mercado del alquiler está creciendo.

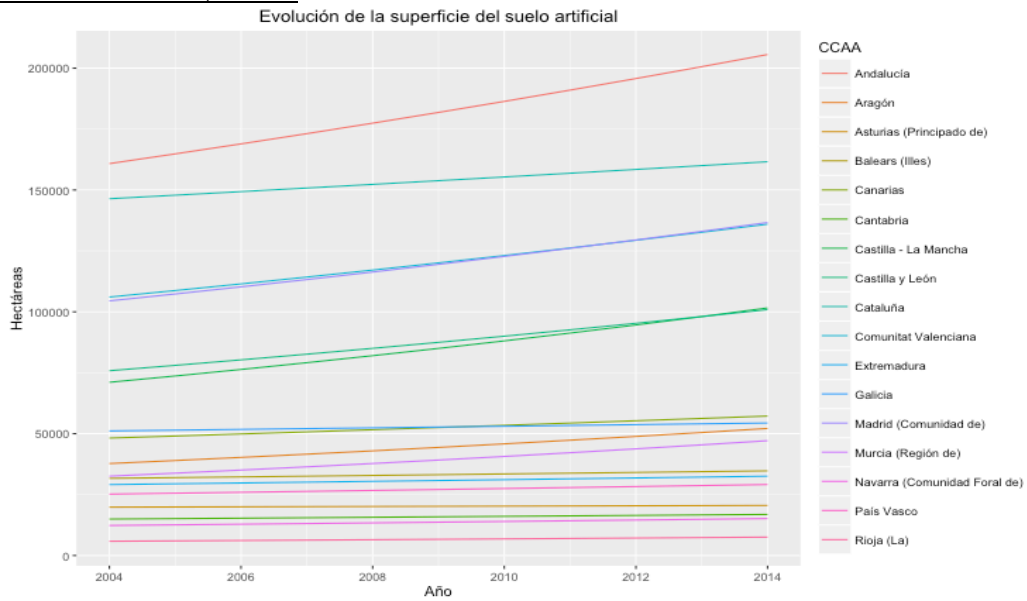
**Figura 3.41:** Gráfico de la evolución del número de viviendas en alquiler por comunidad autónoma entre 2004 y 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

Si se analiza la superficie del suelo artificial, se observa también un crecimiento proporcional en muchas comunidades autónomas, aunque destacan seis comunidades, las cuatro ya mencionadas junto con Castilla - La Mancha y Castilla y León, que tienen una superficie de suelo artificial superior a la mayoría.

**Figura 3.42:** Gráfico de la evolución de la superficie de suelo artificial por comunidad autónoma entre 2004 y 2014.



Fuente: Ministerio de Fomento. Elaboración propia

## IV. TRATAMIENTO DE LAS DESIGUALDADES A TRAVÉS DE ÍNDICES

### 4.1. ¿Qué se entiende por desigualdad? Cómo medirla a través de estadísticos e índices

Para hablar de desigualdad, primero deberíamos definir el concepto “igualdad”. En lo que a este trabajo representa, se puede tomar la tercera acepción de la palabra “igualdad” del diccionario de la Real Academia Española: “Principio que reconoce la equiparación de todos los ciudadanos en derechos y obligaciones”. Por lo tanto, en lo referente a la vivienda, la igualdad representa el poder acceder a una vivienda.

Otras definiciones de igualdad en referencia al acceso a la vivienda podría ser la unión del tratamiento igual a iguales pero garantizando que nadie quede por debajo de un estándar de bienestar.

Hablando desde el punto de vista económico, la igualdad podría ser el poder tener una vivienda independientemente de la situación de la economía, global, regional o personal.

Por lo tanto, se podría definir la desigualdad como el incumplimiento de alguna de estas hipótesis.

Aun así, cuando uno tiene la definición de desigualdad, se plantea otro problema: cómo medirla, interpretarla y compararla.

### 4.2. Definición de estadísticos e índices. Propiedades

Para la definición de los índices se tomará cada variable  $Y$  junto con el vector de  $n$  valores que toma dicha variable,  $Y = \{y_1, y_2, \dots, y_n\}$ . Se define también  $\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$ , como la media de la variable.

Las formas de medir la desigualdad tienen propiedades distintas entre sí, aunque existen un conjunto de propiedades deseables que hacen un estimador de desigualdad preferible a otro según el objeto o objetivo de estudio. Estas propiedades deseables son:

- Independencia de escala: el indicador no debe variar en transformaciones proporcionales de la variable.
- Independencia del tamaño de la población: la desigualdad debe mantenerse si el número de sujetos de la población aumenta de manera proporcional en todos los niveles de la variable.
- Independencia ante cambios de posición: si dos sujetos cambian su posición en la distribución de la variable la desigualdad no debe verse afectada.
- Principio de transferencia:

- Débil: La desigualdad debe disminuir ante la reducción de una observación con valor elevado y el aumento de una observación de valor reducido.
- Fuerte: ante la transferencia de un sujeto con una observación con valor elevado a uno con una observación con valor reducido, la disminución de la desigualdad será mayor cuanto mayor sea la distancia entre ambos.
- Descomposición aditiva: la concentración de una variable para una población debe ser igual a la suma de la desigualdad intragrupal e intergrupala para todos los grupos que la conforman.
- Rango del índice: es interesante que el índice tome valores entre 0 y uno, siendo 0 la representación de la igualdad máxima y 1 la representación de la desigualdad máxima. Esta propiedad permite una mayor comparación entre índices.

Dadas las anteriores propiedades, se presentan algunos estadísticos que miden la desigualdad y que se encuentran entre los más usados hoy en día. Los primeros, son de cálculo sencillo y de posible aplicación en más ámbitos diferentes al análisis de la desigualdad. Estos son el rango, la variancia o el coeficiente de variación.

- Rango: representa la máxima diferencia entre dos valores de una variable.

$$R = y_{Max} - y_{Min}.$$

- Variancia: mide el promedio de la distancia de las observaciones respecto a la media, elevado al cuadrado.

$$V = \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2.$$

- Coeficiente de variación: es una forma de estandarizar la variancia para evitar el efecto de las observaciones extremas y eliminar las unidades de medida.

$$c = \frac{\sqrt{V}}{\bar{y}}.$$

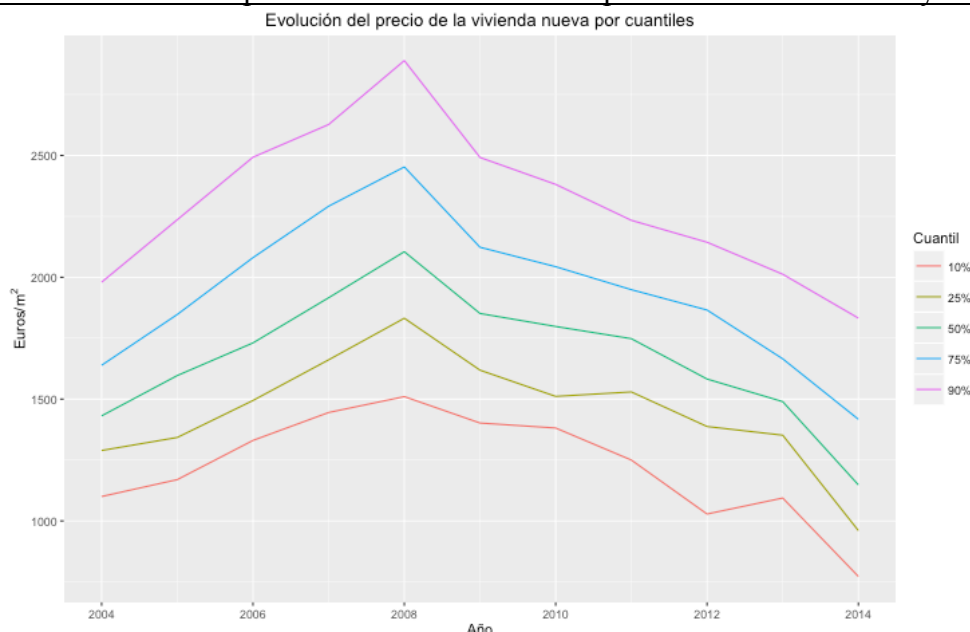
Otra forma de evitar el problema de los valores extremos es mediante la variancia logarítmica, que se define como:

$$v = \frac{1}{n} \sum_{i=1}^n \left[ \log \left( \frac{y_i}{\bar{y}} \right) \right]^2.$$



Otra medida descriptiva interesante son los cuantiles, los cuales son medidas de posición que permiten conocer los puntos característicos de una variable. Si dichos cuantiles se consideran en cada momento temporal, entonces es posible analizar la evolución y dinámica temporal de la distribución de la variable. En este caso, es habitual que se tomen como posiciones características el cuartil 0.1, el cuartil 0.25, el cuartil 0.5 (que equivale a la mediana), el cuartil 0.75 y el cuartil 0.9.

**Figura 4.1:** Evolución del precio de la vivienda nueva por cuantiles entre 2004 y 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

Otra forma de medir la desigualdad es mediante el cálculo de los denominados índices de desigualdad. El más conocido de estos índices es el coeficiente de Gini, y para calcularlo se necesitan ordenar los valores de la variable de menor a mayor, siendo  $y_{(1)}$  el valor mínimo y  $y_{(n)}$  el valor máximo.

$$G = \frac{2}{n^2 \bar{y}} [y_{(1)} + 2y_{(2)} + 3y_{(3)} + \dots + ny_{(n)}] - \frac{n+1}{n}.$$

Este índice es la interpretación numérica de la curva de Lorenz, ya que indica el área de la superficie entre la recta de 45° o línea de equidistribución y la curva de Lorenz. La ventaja de este índice frente a las alternativas anteriores es que representa muy bien dos situaciones. Si todos los valores fueran iguales, el índice tomaría el valor cero, mientras que si hubiera un valor diferente (el más alto) al resto, entonces tomaría el valor uno. La interpretación del índice en el análisis mercado de la vivienda ofrece múltiples posibilidades, por ejemplo en el caso de ver valores de precios medios en áreas territoriales concretas, un índice igual a cero indicaría que el precio es el mismo en todas las áreas, mientras que un

índice igual a uno se daría cuando el precio es cero (la vivienda sería gratuita) en todas las áreas excepto en una.

Otro índice de interés es el Índice de Atkinson. Se define  $\epsilon$  como el parámetro de aversión a la desigualdad, que toma valores no negativos, donde el valor 0 implica indiferencia ante la desigualdad y a medida que este valor aumenta, aumenta la importancia de los valores más extremos; si se hace referencia a rentas, las más pobres. Este parámetro se puede interpretar como una medida de la parte de la renta redistribuida que la sociedad está dispuesta a perder (generalmente en forma de gastos de gestión) como consecuencia de la redistribución. El índice de Atkinson se define una vez fijado el valor del parámetro  $\epsilon$  como:

$$A = 1 - \left[ \frac{1}{n} \sum_{i=1}^n \left[ \frac{y_i}{\bar{y}} \right]^{1-\epsilon} \right]^{\frac{1}{1-\epsilon}}.$$

En el caso que  $\epsilon$  sea igual a 1, el índice de Atkinson se calcula como:

$$A = 1 - \frac{1}{\bar{y}} \left( \prod_{i=1}^n y_i \right)^{\frac{1}{n}}.$$

En la teoría de la información, se quiere valorar la información sobre que un evento  $i$  ocurra. Esta información se representa a través de las probabilidades asociadas a cada uno de los eventos posibles. Pero para una mayor visualización del comportamiento de diferentes eventos, se quiere agregar la información de estos en un solo número. Para ello se define el concepto de entropía.

La entropía define que la probabilidad de que sólo ocurra un evento (que un evento tenga probabilidad 1) no da información, mientras que probabilidades de ocurrencia igual entre todos los eventos representa la máxima información obtenida de unos datos. La función que a partir de la probabilidad da la información de cada evento se define como:

$$h = -\log(p_i).$$

A partir de esta función, se define la entropía como:

$$\text{entropía} = \sum_{i=1}^n p_i h(p_i) = -\sum_{i=1}^n p_i \log(p_i).$$

Con esta fórmula, se obtiene un estadístico que se consigue extraer la información de cada evento  $i$  dándole la importancia que le corresponde dentro del conjunto de eventos, cuya probabilidad respectiva es  $p_i$ .

Theil redefinió la entropía, cambiando la interpretación de  $n$  eventos por  $n$  sujetos de una población y reinterpretando  $p_i$  como las partes asociadas a cada sujeto (*shares*):

$$s_i = \frac{y_i}{n\bar{y}}, \quad \sum_{i=1}^n s_i = 1.$$

A partir de esta definición, se define el Índice de Theil, definido entre 0 (máxima igualdad) y  $\log(n)$  (máxima desigualdad):

$$\begin{aligned} T &= \sum_{i=1}^n \frac{1}{n} h\left(\frac{1}{n}\right) - \sum_{i=1}^n s_i h(s_i) = \sum_{i=1}^n s_i \left[ h\left(\frac{1}{n}\right) - h(s_i) \right] = \sum_{i=1}^n s_i \left[ \log\left(\frac{1}{n}\right) - \log(s_i) \right] \\ &= \frac{1}{n} \sum_{i=1}^n \frac{y_i}{\bar{y}} \log\left(\frac{y_i}{\bar{y}}\right). \end{aligned}$$

Se puede obtener el índice de Theil Normalizado, que toma valores entre 0 y 1:

$$T_n = \frac{T}{\log(n)}.$$

Kolm redefinió el índice de Atkison, porque este tenía la propiedad de independencia de escala y, por lo tanto, ante la situación de dos rentas de diferente valor que crecen proporcionalmente (se multiplican por un mismo valor) el índice de Atkinson se mantiene constante, aunque la diferencia absoluta entre los sujetos aumente. Ante esa situación, Kolm definió un índice que se adaptaba ante esta variación, así el valor del índice se multiplicaba por el mismo valor que la variable. Por el contrario, si se añade el mismo valor a todos los sujetos, el índice de Atkison varía, mientras que el de Kolm se mantiene constante.

$$K = \frac{1}{\epsilon} \ln \left[ \frac{1}{n} \sum_{i=1}^n e^{\epsilon(\bar{y}-y_i)} \right].$$

Finalmente, una medida de desigualdad que cumple con las propiedades de independencia de escala, principio de transferencia débil y descomposición aditiva es el índice de entropía generalizado, definido por:

$$E_\theta = \frac{1}{\theta^2 - \theta} \left[ \frac{1}{n} \sum_{i=1}^n \left[ \frac{y_i}{\bar{y}} \right]^\theta - 1 \right].$$

#### 4.3. Aplicación de los métodos de medición presentados a los datos de estudio

Uno de los ámbitos que más se ha beneficiado de las diferentes formas de medir la desigualdad es el análisis de la renta per cápita de una población. La definición de los índices de desigualdad permite una fácil comparación entre países o bien un análisis longitudinal de

cómo evolucionan las disparidades dentro de un mismo país. Para medir la desigualdad, la mayoría de estudios que estudian renta per cápita de una población, persiguen una fácil interpretación de los índices. Sin embargo, el análisis de la desigualdad y la implementación de los índices pueden aplicarse también a otras variables como las de estudio en este trabajo, a la vez que pueden utilizarse los índices para calcular la dispersión dentro de las variables.

La primera de estas variables a analizar mediante los estadísticos descriptivos es el precio de la vivienda nueva en España, siendo los distintos sujetos cada una de las áreas urbanas de estudio. Observando el rango, de 2577.60 euros el metro cuadrado, se puede observar una fuerte diferencia entre el precio de vivienda nueva más elevado y más bajo. También se observa una alta variancia, pero teniendo en cuenta el coeficiente de variación de valor 0.35, que elimina el efecto de las unidades, se puede deducir que no hay una fuerte diferencia de precios entre la mayoría de las áreas urbanas de España.

**Tabla 4.1:** Estadísticos descriptivos para medir la desigualdad de las variables de estudio para el año 2014.

	<b>Rango</b>	<b>Variancia</b>	<b>Coef. Var</b>
<i>Población</i>	31954590.00	12341244722128.13	4.69
<i>Densidad de población</i>	6247.10	805002.89	1.17
<i>Extranjeros</i>	819413.00	8115638762.85	2.95
<i>Edad media</i>	10.80	3.84	0.05
<i>Precio vivienda nueva</i>	2577.60	190595.19	0.35
<i>Precio vivienda usada</i>	1868.00	113459.67	0.29
<i>Tasación vivienda nueva</i>	2516.10	180343.60	0.31
<i>Tasación vivienda usada</i>	2133.90	119641.42	0.30
<i>Transacciones anuales</i>	49253.00	42791654.80	2.17
<i>Densidad viviendas familiares</i>	304.25	2656.55	0.92
<i>Densidad viviendas familiares principales</i>	2485479.53	124920956002.83	2.25
<i>Densidad viviendas secundarias</i>	1181903.02	42466855018.74	1.71
<i>Densidad viviendas vacías</i>	261516.16	1808406133.10	1.63
<i>Viviendas familiares principales</i>	4600.30	373404.24	3.59
<i>Viviendas secundarias</i>	1283374.57	32562188178.64	2.37
<i>Viviendas vacías</i>	12357848.51	3061792498923.62	3.76
<i>Otras viviendas</i>	80340.62	155931414.53	1.81
<i>Propiedad pagada</i>	214153.92	830272729.78	2.84
<i>Pagos pendientes</i>	14120.05	3519174.85	1.65
<i>Propiedad por herencia o donación</i>	41537.05	39621191.56	1.91
<i>Alquiler</i>	110527.07	217121038.85	1.92
<i>Gratuita</i>	66.23	211.57	0.51
<i>Otro régimen</i>	285.35	1073.16	0.90

Fuente: Elaboración propia.

Observando la Tabla 4.1 se puede ver como el mayor coeficiente de variación es el correspondiente a la variable “Población”, lo que indica que las áreas urbanas analizadas son muy diferentes entre sí porque el número de personas que residen en ellas es muy dispar. La siguiente variable en orden decreciente de coeficiente de variación es la de “Viviendas vacías”. Ello indicaría grandes diferencias en el número de viviendas vacías por áreas urbanas en el año 2014. También es interesante observar el coeficiente de variación de la variable “Transacciones anuales” que indica grandes diferencias entre el número de transacciones anuales entre diferentes regiones.

Si se analizan las desigualdades del precio de la vivienda nueva a través de los índices anteriormente presentados, el primero, el índice de Gini, muy cercano a 0, de valor 0.18, nos indica que no hay mucha desigualdad entre los precios de la vivienda nueva entre las distintas áreas urbanas.

El índice de Atkinson, por su parte, muy cercano a 0, nos indica que el valor del precio de la vivienda nueva está bastante concentrado alrededor de un mismo nivel, con excepción de alguna área urbana que tiene un precio de la vivienda nueva elevado. Por lo tanto, el nivel medio de precios pagado por la mayoría de la población es parecido, lo que representaría un nivel parecido de bienestar.

Lo mismo se puede observar con el Índice de Theil Normalizado, muy cercano al 0, el cual indica poca diferencia entre los sujetos de estudio respecto a la variable precio de la vivienda nueva.

De nuevo una de las variables que denota más desigualdad entre áreas urbanas es la “Vivienda vacía”.

Finalmente, si se mide la desigualdad a través del índice de entropía generalizada, se observa que la mayor desigualdad se encuentra en el número de viviendas vacías por área urbana, seguida por la población y el número de viviendas familiares principales, las cuales están relacionadas. A más población, mayor número de viviendas familiares principales.

Destaca también la densidad de viviendas secundarias.

**Tabla 4.2:** Índices para medir la desigualdad de las variables de estudio para el año 2014.

	<i>Gini</i>	<i>Atkinson</i>	<i>Theil</i>	<i>Kolm</i>	<i>Entropía Generalizada</i>
<i>Población</i>	0.81	0.59	0.31	Inf	1.37
<i>Densidad de población</i>	0.48	0.19	0.10	698.51	0.43
<i>Extranjeros</i>	0.66	0.38	0.19	Inf	0.82
<i>Edad media</i>	0.03	0.00	0.00	1.02	0.00
<i>Precio vivienda nueva</i>	0.19	0.03	0.01	791.10	0.06
<i>Precio vivienda usada</i>	0.15	0.02	0.01	549.64	0.04
<i>Tasación vivienda nueva</i>	0.16	0.02	0.01	705.66	0.04
<i>Tasación vivienda usada</i>	0.15	0.02	0.01	566.78	0.04
<i>Transacciones anuales</i>	0.64	0.35	0.17	Inf	0.77
<i>Densidad viviendas familiares</i>	0.44	0.16	0.09	46.22	0.38
<i>Densidad viviendas familiares principales</i>	0.65	0.35	0.17	Inf	0.75
<i>Densidad viviendas secundarias</i>	0.69	0.40	0.26	Inf	1.14
<i>Densidad viviendas vacías</i>	0.58	0.28	0.13	Inf	0.59
<i>Viviendas familiares principales</i>	0.80	0.57	0.34	161.57	1.50
<i>Viviendas secundarias</i>	0.69	0.40	0.20	Inf	0.89
<i>Viviendas vacías</i>	0.83	0.60	0.36	Inf	1.59
<i>Otras viviendas</i>	0.67	0.36	0.20	Inf	0.88
<i>Propiedad pagada</i>	0.71	0.44	0.23	Inf	1.02
<i>Pagos pendientes</i>	0.58	0.27	0.13	993.37	0.59
<i>Propiedad por herencia o donación</i>	0.66	0.36	0.19	Inf	0.86
<i>Alquiler</i>	0.59	0.30	0.14	Inf	0.63
<i>Gratuita</i>	0.29	0.07	0.03	16.01	0.14
<i>Otro régimen</i>	0.35	0.12	Inf	27.90	Inf

Fuente: Elaboración propia.

Si se comparan los estadísticos descriptivos e índices obtenidos con datos del año 2014 con los estadísticos descriptivos e índices obtenidos con datos del año 2004, se puede observar la evolución temporal de las variables respecto a las distintas áreas urbanas, donde destaca la disminución de la desigualdad en el valor de las tasaciones de vivienda nueva y el aumento de la desigualdad en el número de viviendas familiares principales y viviendas vacías.

**Tabla 4.3:** Estadísticos descriptivos para medir la desigualdad de las variables de estudio para el año 2004.

	<b>Rango</b>	<b>Variancia</b>	<b>Coef. Var</b>
<i>Población</i>	5439807.00	611452496389.04	2.27
<i>Densidad de población</i>	5021.70	594602.05	1.13
<i>Extranjeros</i>	628384.00	6727240620.51	3.07
<i>Edad media</i>	10.50	5.36	0.06
<i>Precio vivienda nueva</i>	1898.70	146827.29	0.25
<i>Precio vivienda usada</i>	1862.20	153363.05	0.29
<i>Tasación vivienda nueva</i>	100932.00	203841117.51	2.06
<i>Tasación vivienda usada</i>	4061.65	656123.36	0.45
<i>Transacciones anuales</i>	3046.16	314206.54	0.40
<i>Densidad viviendas familiares</i>	254.31	2081.82	0.92
<i>Densidad viviendas familiares principales</i>	1908123.20	77896261157.69	2.31
<i>Densidad viviendas secundarias</i>	613.45	11664.46	0.35
<i>Densidad viviendas vacías</i>	188546.74	898400176.49	1.61
<i>Viviendas familiares principales</i>	37.07	71.18	0.50
<i>Viviendas secundarias</i>	394007.80	3532238350.81	2.06
<i>Viviendas vacías</i>	236.53	1934.08	0.36
<i>Otras viviendas</i>	24478.29	17006216.94	1.53
<i>Propiedad pagada</i>	606320.86	8039491962.45	2.42
<i>Pagos pendientes</i>	222447.44	982803900.27	2.28
<i>Propiedad por herencia o donación</i>	48719.94	57546771.22	1.77
<i>Alquiler</i>	82383.94	140306666.40	1.90
<i>Gratuita</i>	67.58	233.21	0.48
<i>Otro régimen</i>	85.13	330.17	0.56

Fuente: Elaboración propia.

**Tabla 4.4:** Índices para medir la desigualdad de las variables de estudio para el año 2004.

	<i>Gini</i>	<i>Atkinson</i>	<i>Theil</i>	<i>Kolm</i>	<i>Entropía Generalizada</i>
<i>Población</i>	0.65	0.36	0.17	Inf	0.77
<i>Densidad de población</i>	0.48	0.19	0.10	622.39	0.43
<i>Extranjeros</i>	0.74	0.48	0.27	Inf	1.19
<i>Edad media</i>	0.03	0.00	0.00	1.46	0.00
<i>Precio vivienda nueva</i>	0.14	0.02	0.01	676.81	0.03
<i>Precio vivienda usada</i>	0.15	0.02	0.01	693.01	0.04
<i>Tasación vivienda nueva</i>	0.62	0.33	0.16	Inf	0.72
<i>Tasación vivienda usada</i>	0.24	0.05	0.02	1232.89	0.09
<i>Transacciones anuales</i>	0.22	0.04	0.02	1046.04	0.08
<i>Densidad viviendas familiares</i>	0.44	0.16	0.08	39.84	0.38
<i>Densidad viviendas familiares principales</i>	0.66	0.37	0.18	Inf	0.79
<i>Densidad viviendas secundarias</i>	0.18	0.03	0.01	192.48	0.06
<i>Densidad viviendas vacías</i>	0.59	0.29	0.15	Inf	0.66
<i>Viviendas familiares principales</i>	0.27	0.06	0.03	7.70	0.13
<i>Viviendas secundarias</i>	0.62	0.32	0.15	Inf	0.68
<i>Viviendas vacías</i>	0.18	0.03	0.01	47.11	0.05
<i>Otras viviendas</i>	0.62	0.32	0.18	Inf	0.80
<i>Propiedad pagada</i>	0.67	0.38	0.19	Inf	0.84
<i>Pagos pendientes</i>	0.65	0.36	0.17	Inf	0.77
<i>Propiedad por herencia o donación</i>	0.64	0.33	0.17	Inf	0.77
<i>Alquiler</i>	0.60	0.30	0.14	Inf	0.64
<i>Gratuita</i>	0.27	0.06	0.03	17.43	0.12
<i>Otro régimen</i>	0.31	0.09	Inf	24.13	Inf

Fuente: Elaboración propia.

Dadas las diferencias entre población entre áreas urbanas, se analiza la dispersión de algunas de las variables por cápita para contrastar diferencias observadas en valor absoluto y ver el efecto de la población.



**Tabla 4.5:** Comparativa de los estadísticos descriptivos para medir la desigualdad de las variables per cápita de estudio entre 2004 y 2014.

	<b>Rango 2004</b>	<b>Rango 2014</b>	<b>Variancia 2004</b>	<b>Variancia 2014</b>	<b>Coef. Var 2004</b>	<b>Coef. Var 2014</b>
<i>Transacciones anuales</i>	0.07	0.04	0.00	0.00	0.92	0.74
<i>Viviendas familiares principales</i>	0.26	0.52	0.00	0.00	0.09	0.13
<i>Viviendas secundarias</i>	0.01	7.81	0.00	0.96	0.60	1.99
<i>Viviendas vacías</i>	0.55	0.23	0.01	0.00	0.99	0.49
<i>Otras viviendas</i>	0.00	0.00	0.00	0.00	0.88	0.76
<i>Propiedad pagada</i>	0.22	0.37	0.00	0.00	0.29	0.33
<i>Pagos pendientes</i>	0.00	2.11	0.00	0.12	0.61	0.57
<i>Propiedad por herencia o donación</i>	0.09	0.09	0.00	0.00	1.11	0.88
<i>Alquiler</i>	0.06	0.06	0.00	0.00	0.13	0.53
<i>Gratuita</i>	0.03	0.01	0.00	0.00	0.17	0.51
<i>Otro régimen</i>	0.03	0.03	0.00	0.00	0.50	0.72

Fuente: Elaboración propia.

**Tabla 4.6:** Comparativa de los índices de Gini, Atkinson y Theil de las variables per cápita de estudio entre 2004 y 2014.

	<b>Gini 2004</b>	<b>Gini 2014</b>	<b>Atkinson 2004</b>	<b>Atkinson 2014</b>	<b>Theil 2004</b>	<b>Theil 2014</b>
<i>Transacciones anuales</i>	0.46	0.30	0.18	0.09	0.10	0.04
<i>Viviendas familiares principales</i>	0.04	0.05	0.00	0.01	0.00	0.01
<i>Viviendas secundarias</i>	0.34	0.65	0.10	0.35	0.05	0.20
<i>Viviendas vacías</i>	0.40	0.27	0.14	0.06	0.06	0.04
<i>Otras viviendas</i>	0.45	0.41	0.16	0.14	0.09	0.08
<i>Propiedad pagada</i>	0.14	0.18	0.02	0.03	0.01	0.03
<i>Pagos pendientes</i>	0.34	0.29	0.10	0.07	0.06	0.04
<i>Propiedad por herencia o donación</i>	0.42	0.42	0.15	0.14	0.07	0.07
<i>Alquiler</i>	0.08	0.28	0.00	0.07	0.00	0.04
<i>Gratuita</i>	0.09	0.27	0.01	0.06	0.00	0.04
<i>Otro régimen</i>	0.26	0.35	0.05	0.11	0.02	0.06

Fuente: Elaboración propia.

**Tabla 4.7:** Comparativa de los índices de Kolm y la entropía generalizada de las variables per cápita de estudio entre 2004 y 2014.

	<b>Kolm 2004</b>	<b>Kolm 2014</b>	<b>Entropía Generalizada 2004</b>	<b>Entropía Generalizada 2014</b>
<i>Transacciones anuales</i>	0.00	0.00	0.45	0.19
<i>Viviendas familiares principales</i>	0.00	0.00	0.00	0.05
<i>Viviendas secundarias</i>	0.00	0.13	0.24	0.88
<i>Viviendas vacías</i>	0.00	0.00	0.28	0.16
<i>Otras viviendas</i>	0.00	0.00	0.38	0.36
<i>Propiedad pagada</i>	0.00	0.00	0.03	0.11
<i>Pagos pendientes</i>	0.00	0.03	0.25	0.20
<i>Propiedad por herencia o donación</i>	0.00	0.00	0.31	0.33
<i>Alquiler</i>	0.00	0.00	0.01	0.17
<i>Gratuita</i>	0.00	0.00	0.02	0.16
<i>Otro régimen</i>	0.00	0.00	0.11	0.26

Fuente: Elaboración propia.

Aquí, la desigualdad en el número de viviendas vacías por cápita disminuye entre 2004 y 2014, mientras que sí aumenta el número de viviendas secundarias per cápita, es decir, el número de viviendas secundarias per cápita aumenta más en algunas áreas urbanas respecto a otras, doblando la desigualdad según los índices de Gini, Theil y Atkinson. El rango de valor 7.81 en el año 2014 respecto al rango de valor 0.01 del año 2004 demuestra la importancia del mercado de vivienda secundaria en algunas áreas urbanas.

## 5. MODELO DE PANEL PARA EXPLICAR EL COMPORTAMIENTO DEL PRECIO DE LA VIVIENDA NUEVA

El objetivo de estudio es ver la evolución del precio de la vivienda nueva en España en las distintas áreas urbanas a lo largo del periodo 2004-2014. Para ello, se usará un modelo de regresión para datos de panel con efectos fijos, dado que los objetos de estudio son constantes y mantienen la mayoría de sus respectivas características.

El modelo de panel permite analizar el efecto de las variables explicativas sobre la variable endógena en un momento dado y a lo largo del tiempo, lo que permite hacer un análisis dinámico.

### 5.1. Definición del modelo

Las unidades de observación (o sujetos), las áreas urbanas, se diferencian por el índice  $i$ , que toma valores de 1 a  $n$ , con  $n=86$  para cada una de las distintas áreas urbanas. Por su parte, observamos  $T$  momentos temporales distintos, representados en años con el índice  $t$  para cada uno de los años de estudio (de 2004 a 2014).

Definidos los sujetos y el período temporal, se define como  $y_{it}$  la respuesta del sujeto  $i$  en el momento temporal  $t$  y el conjunto  $x_{it}$  como las variables explicativas para el sujeto  $i$  en el momento  $t$ . En este caso concreto, se define como variable respuesta el *Precio de la vivienda nueva* y como conjunto de variables explicativas las  $K$  variables restantes presentadas anteriormente. Así, definimos  $x_{it} = \{x_{it,1}, x_{it,2}, \dots, x_{it,K}\}'$  como el vector de variables explicativas para el individuo  $i$  en el momento  $t$ .

Por lo tanto, los datos para cada individuo  $i$  están definidos como:

$$\{x_{it}, y_{it}\}.$$

Con los datos, se puede definir el modelo de regresión básico para poder analizar la relación entre la variable respuesta y las variables explicativas en un momento dado del tiempo  $t$ :

$$y_{it} = \alpha + \beta_1 x_{it,1} + \beta_2 x_{it,2} + \dots + \beta_K x_{it,K} + \epsilon_{it},$$

donde  $\alpha, \beta_1, \dots, \beta_K$  son parámetros que definen como afectan las variables explicativas sobre la variable respuesta y  $\epsilon_{it}$  es un término de perturbación aleatoria con esperanza igual a cero.

### 5.1.1. Supuestos del modelo

De cara a la inferencia, se deben asumir cuatro supuestos sobre los datos:

- El primero, la forma lineal del modelo definido anteriormente.
- El segundo, el conjunto de variables explicativas,  $\{x_{it,1}, x_{it,2}, \dots, x_{it,K}\}$ , deben ser variables no estocásticas, no aleatorias.
- La variancia de  $y_{it}$  es constante ( $\text{Var } y_{it} = \sigma$ ).
- La variable respuesta,  $y_{it}$ , son variables aleatorias independientes, no se elige una parte de concreta de la población, así no se sesgan las estimaciones. Por lo tanto, se define que no hay correlación entre  $y_{it}$  a lo largo del tiempo ni entre los diferentes sujetos.

Para este estudio, el modelo que interesa es el de efectos fijos, el cual asume los supuestos anteriormente definidos, modificando el primero, el cual redefine como:

$$E(y_{it}) = \alpha_i + \beta_1 x_{it,1} + \beta_2 x_{it,2} + \dots + \beta_K x_{it,K} = \alpha_i + x'_{it}\beta,$$

es decir, se mantienen los parámetros poblacionales,  $\beta$ , que definen la aportación de cada variable explicativa sobre la esperanza de la variable respuesta, pero varía el parámetro  $\alpha$ , en este modelo definido como  $\alpha_i$ , ya que define la aportación de un sujeto  $i$  concreto sobre el valor esperado de la variable respuesta. Este último parámetro nos ayuda a controlar la heterogeneidad entre las unidades de observación, siendo un parámetro que se estima a posteriori, cuando se tienen varias observaciones. Además, este parámetro en el modelo de panel es una ventaja respecto a los modelos de regresión en un único período temporal, ya que en este último caso el efecto de  $\alpha_i$  se encuentra mezclado con el término de perturbación, representado por el residuo  $\epsilon_i$ , que representa el error entre el valor estimado y el valor real. Por lo tanto, en el modelo de panel, al incluir el parámetro  $\alpha_i$  se consigue reducir la variancia del término de perturbación y, en consecuencia, reducir la variabilidad del modelo y conseguir estimaciones más eficientes.

### 5.1.2. Estimación de los parámetros poblacionales.

Dada la ecuación:

$$E(y_{it}) = \alpha_i + \beta_1 x_{it,1} + \beta_2 x_{it,2} + \dots + \beta_K x_{it,K} = \alpha_i + x'_{it}\beta.$$

Según el Teorema de Gauss-Markov, los mejores estimadores lineales sin sesgo de los parámetros  $\beta$  son los estimados por mínimos cuadrados ordinarios, definidos por:

$$b = \left( \sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' \right)^{-1} \left( \sum_{i=1}^n \sum_{t=1}^T (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i) \right),$$

donde  $b = (b_1, b_2, \dots, b_K)'$  y, seguidamente, se puede estimar  $a_i$  como:

$$a_i = \bar{y}_i - \bar{x}_i' b.$$

Definimos la matriz de pesos (*weights*) para cada sujeto como:

$$W_i = \sum_{t=1}^T (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)'$$

Los estimadores  $b$  y  $a_i$  son estimadores insesgados, y además, tienen mínima variancia, definida como:

$$Var b = \sigma^2 \left( \sum_{i=1}^n W_i \right)^{-1}.$$

La estimación del parámetro  $\sigma^2$  se hace a través del método del análisis de la variancia, ANOVA. Se define  $e_{it} = y_{it} + a_i + x_{it}' b$  como los residuos del modelo, es decir, la diferencia entre los valores observados y estimados. La suma de los residuos al cuadrado se denota como  $Error SS = \sum_{it} e_{it}^2$ . A partir de este valor, se estima  $\sigma^2$  a través de  $s^2$ :

$$s^2 = \frac{Error SS}{N - (n + K)} = Error MS,$$

donde  $s$  es la desviación estándar residual y  $N$  es el total de observaciones, es decir,  $nT$  en el caso de un panel de  $n$  observaciones a lo largo de  $T$  periodos.

### 5.1.3. Bondad del ajuste.

Finalmente, se obtiene el coeficiente de determinación,  $R^2$ , que toma valores entre 0 y 1 (siendo 1 el valor del ajuste perfecto), que sirve para determinar la bondad de ajuste del modelo. Para ello, se define la suma de cuadrados del total,  $Total SS = \sum_{it} (y_{it} - \bar{y})^2$ , y la suma de cuadrados de la regresión,  $Regression SS = \sum_{it} (a_i + x_{it}' b - \bar{y})^2$ . Con esto, se define  $R^2$  como:

$$R^2 = \frac{Regression SS}{Total SS}.$$

También se puede determinar el coeficiente de determinación ajustado por los grados de libertad:

$$R_{adj}^2 = \frac{(Error SS)/(N - (n + K))}{(Total SS)/(N - 1)}.$$

### 5.1.4. Especificaciones del modelo y diagnósticos.

Para saber si el parámetro  $\alpha$  tiene diferente aportación según el sujeto, se debe realizar el test de heterogeneidad, y, por lo tanto, si tiene sentido realizar el modelo de panel o no.

Para ello, se determina la hipótesis nula de homogeneidad como  $H_0: \alpha_1 = \alpha_2 = \dots = \alpha_K = \alpha$ , siguiendo un test de la F.

Para realizar el test, se deben estimar el modelo completo ( $E(y_{it}) = \alpha_i + x'_{it}\beta$ ) y reducido ( $E(y_{it}) = \alpha + x'_{it}\beta$ ) y obtener el *Error SS* y  $s^2$  del primero y el *Error SS<sub>reduced</sub>* del segundo y calcular el estimador *F-ratio*  $= \frac{\text{Error SS}_{reduced} - \text{Error SS}}{(n-1)s^2} \sim F_{n-1, N-(n+K)}$ . Si el valor del *F-ratio* es mayor al valor de  $F_{n-1, N-(n+K)}$ , se rechaza la hipótesis nula y, por lo tanto, tiene sentido realizar un modelo de panel con los datos dados, donde los valores  $\alpha_1, \alpha_2, \dots, \alpha_K$  toman valores distintos y, por lo tanto, disminuyen la variancia del modelo y en consecuencia el error de éste.

Otro parámetro que puede interesar conocer es la influencia de una observación en el modelo. Para ellos, se pueden realizar análisis de sensibilidad, como el cálculo de la *distancia de Cook*. Aun así, en los modelos de panel de efectos fijos, parte de la influencia de una observación queda reflejada en el parámetro específico para cada sujeto,  $\alpha_i$ .

Otra forma de calcular la influencia de un sujeto es calcular su impacto sobre los coeficientes de la regresión a través del estadístico *B*:

$$B_i(b) = \frac{(b - b_{(i)})' (\sum_{i=1}^n W_i) (b - b_{(i)})}{K} \sim \chi^2,$$

donde  $b_{(i)}$  es la estimación de  $b$  por mínimos cuadrados ordinarios omitiendo al sujeto  $i$ , y  $B_i(b)$  mide la distancia entre los coeficientes calculados con y sin el sujeto  $i$ , lo que permite conocer el impacto de un sujeto  $i$  determinado. El estadístico sigue una distribución Chi-cuadrado con un grado de libertad bajo la hipótesis nula (no influencia).

La diferencia entre este estimador y la distancia de Cook es que el primero calcula el efecto por sujeto y la segunda el efecto por observación.

Para la estimación de modelos de regresión, un supuesto importante es el de homoscedasticidad, es decir, que la variabilidad sea constante entre todas las observaciones. Si esta variabilidad no fuera constante, la eficiencia de los coeficientes de la regresión disminuiría, pasando a realizar estimaciones sesgadas. En el caso de datos de panel, la variabilidad debe ser controlada entre los sujetos,  $i$ , y a lo largo del tiempo,  $t$ .

En caso de presencia de heteroscedasticidad, hay distintas técnicas para tratar este "problema". La forma más sencilla, es la transformación del tratamiento de la variable respuesta.

Para detectar la presencia de heteroscedasticidad también existen distintas técnicas. La primera, es la representación gráfica de los residuos frente a los valores ajustados, a fin de detectar patrones de comportamiento. Otra forma de detectar la heteroscedasticidad es mediante el test de Breusch-Pagan. La hipótesis nula asociada al test de Breusch-Pagan es  $H_0: \text{Var } y_{it} = \sigma^2$ , es decir, presencia de homocedasticidad en el modelo especificado.

Otra propiedad importante para la validación del modelo es el supuesto de consistencia de los estimadores y no omisión de variables relevantes. Este último supuesto se valida mediante el test de Hausman, el cual compara los estimadores de efectos aleatorios frente a los de efectos fijos. En la hipótesis nula, ambos son iguales y se prefiere el estimador de efectos aleatorios, y en la alternativa se preferirá el estimador de efectos fijos por ser insesgado y consistente aunque se omitan variables. El estadístico del test de Hausman se define como:

$$H = (\beta_c - \beta_e)' / (V_c - V_e)^{-1} (\beta_c - \beta_e) \sim \chi_k^2$$

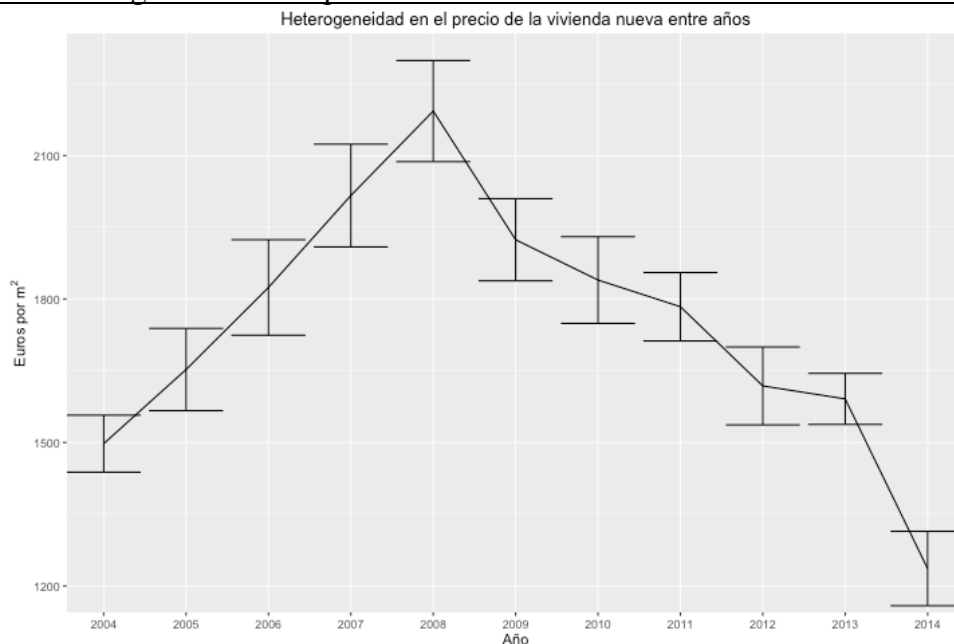
donde  $\beta_c$  es el vector de estimaciones de efectos fijos,  $V_c$  es la matriz de covarianzas de las estimaciones de efectos fijos,  $\beta_e$  es el vector de estimaciones de efectos aleatorios y, finalmente,  $V_e$  es la matriz de covarianzas de las estimaciones de efectos aleatorios. Por su parte, los grados de libertad,  $k$ , son el número de variables incluidas en los efectos fijos.

Previo al análisis a través del modelo de regresión, se realiza un análisis exploratorio de los datos, para ver el comportamiento de la variable respuesta, así como de las variables explicativas y la relación entre ellas.

Con este análisis, ya se pueden obtener conclusiones sobre los datos. Además, el análisis exploratorio permite detectar observaciones o sujetos atípicos.

#### 5.1.4.1. Análisis de la variable respuesta en el estudio del precio de la vivienda.

**Figura 5.1: Heterogeneidad en el precio de la vivienda en áreas urbanas entre 2004 y 2014.**



Fuente: Ministerio de Fomento. Elaboración propia.

El gráfico anterior muestra la media de los precios de la vivienda nueva juntamente con su intervalo de confianza al 95% a lo largo del período temporal estudiado.

Se observa un comportamiento creciente, constante, en los precios de la vivienda nueva entre los años 2004 y 2008, primer año de la crisis actual. A partir de este año, los precios empiezan a descender, con su caída más pronunciada entre 2013 y 2014.

La variabilidad observada entre los 86 sujetos de estudio no varía mucho a lo largo de los 11 años de estudio.

También es interesante ver la cantidad de valores extremos o *outliers* y la distancia de estos con las demás observaciones.

**Figura 5.2:** Representación de *outliers* en el precio de la vivienda nueva según área urbana entre 2004 y 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

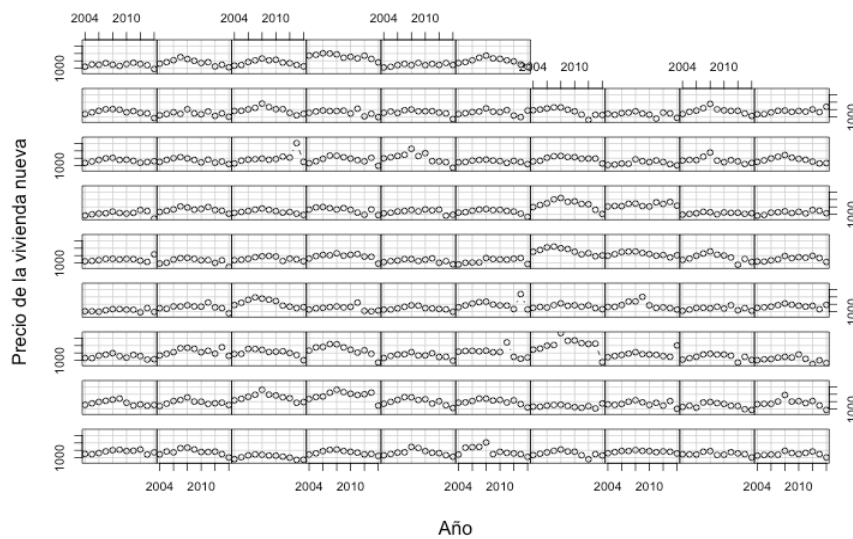
Destaca la no presencia de *outliers* en el año 2007 y como, a partir del año siguiente, 2008, la distancia de los *outliers* es mayor a la distancia a los años previos a 2007.

Es importante conocer si los valores atípicos se concentran en unos mismos sujetos o son valores extremos de un sujeto y momento temporal concreto, aunque en este trabajo los valores extremos son debidos básicamente a la diferencia de las características de las distintas regiones.

Con un *trellis plot* se puede observar el comportamiento de cada uno de los sujetos para la variable:



**Figura 5.3:** Evolución del precio de la vivienda nueva por área urbana entre 2004 y 2014.



Fuente: Ministerio de Fomento. Elaboración propia.

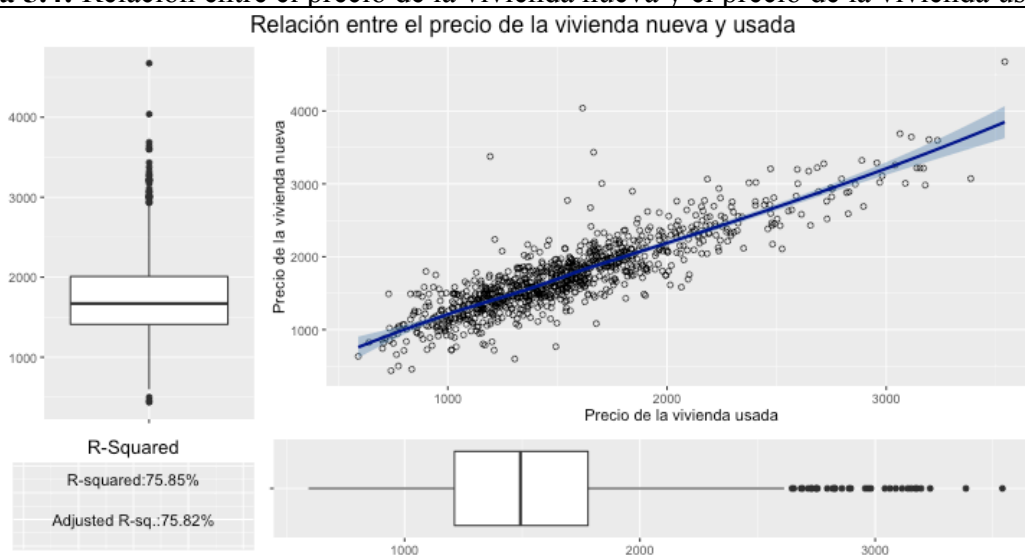
No se pueden observar diferencias aparentes respecto al comportamiento definido anteriormente, pero sí se puede observar la presencia de algún comportamiento atípico en algún momento del tiempo. Por ejemplo, hay algunas áreas urbanas que tienen valores muy alejados del resto en un momento temporal concreto, en general en la segunda mitad del periodo analizado.

#### 5.1.4.2. Análisis de variables explicativas.

Se presenta un análisis exploratorio para la relación entre las variables explicativas del modelo y la variable respuesta.

La primera variable explicativa presentada es el precio de la vivienda usada:

**Figura 5.4:** Relación entre el precio de la vivienda nueva y el precio de la vivienda usada.



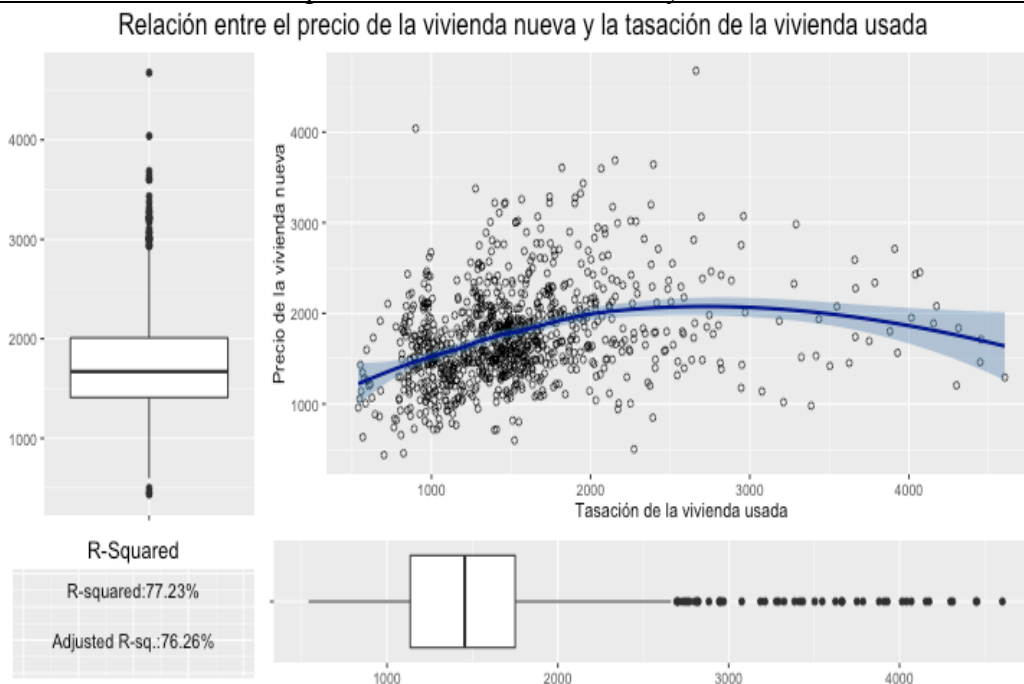
Fuente: Ministerio de Fomento. Elaboración propia.

En la Figura 5.4 se observa claramente la relación lineal entre estas dos variables, siendo la línea azul continua la media del precio de la vivienda usada en relación al precio de la vivienda nueva y la sombra azul colindante marca el intervalo de confianza del 95% respecto a esta media. La variable “precio de la vivienda usada” explica un 75% de la variancia sistemática de la variable “precio de la vivienda nueva”.

La correlación lineal (coeficiente de correlación lineal de Pearson) entre ambas variables es de 0.8709.

La segunda variable presentada es la tasación de la vivienda usada:

**Figura 5.5:** Relación entre el precio de la vivienda nueva y la tasación de la vivienda usada.



Fuente: Ministerio de Fomento. Elaboración propia.

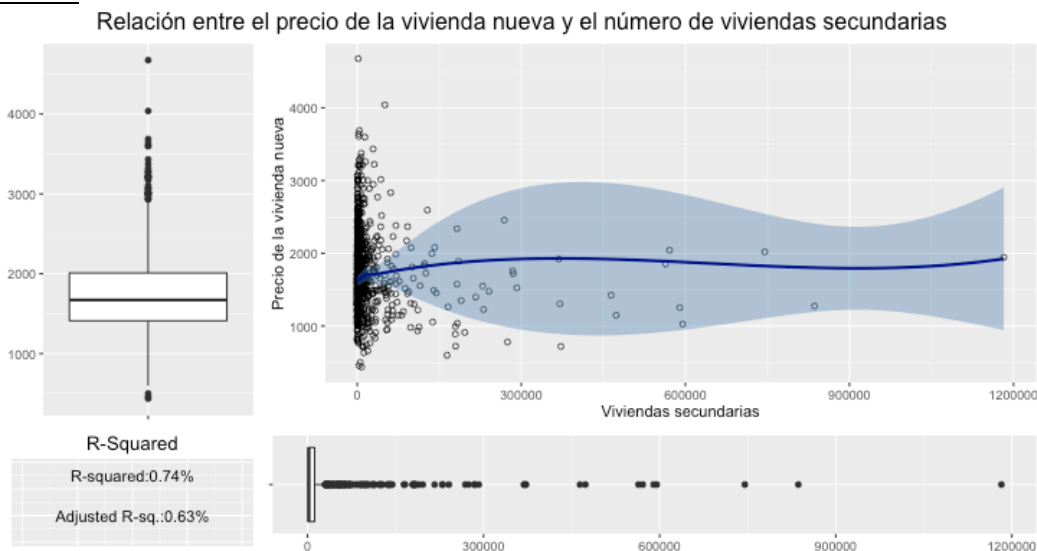
En la Figura 5.5 no se observa la clara relación lineal que se apreciaba entre la variable explicativa anterior y la variable respuesta. Aunque visualmente esta relación no sea tan explícita, la variable “tasación de la vivienda usada” explica un 77.23% de la variancia sistemática de la variable respuesta.

En este caso, la mayoría de observaciones se encuentran en el recuadro donde los precios y las tasaciones se encuentran entre 1000 y 2000 euros. Se observa también una mayor amplitud en el intervalo de confianza respecto al anterior caso.

La correlación lineal entre las variables es de 0.2779.

La tercera variable presentada es el número de viviendas secundarias:

**Figura 5.6:** Relación entre el precio de la vivienda nueva y el número de viviendas secundarias.



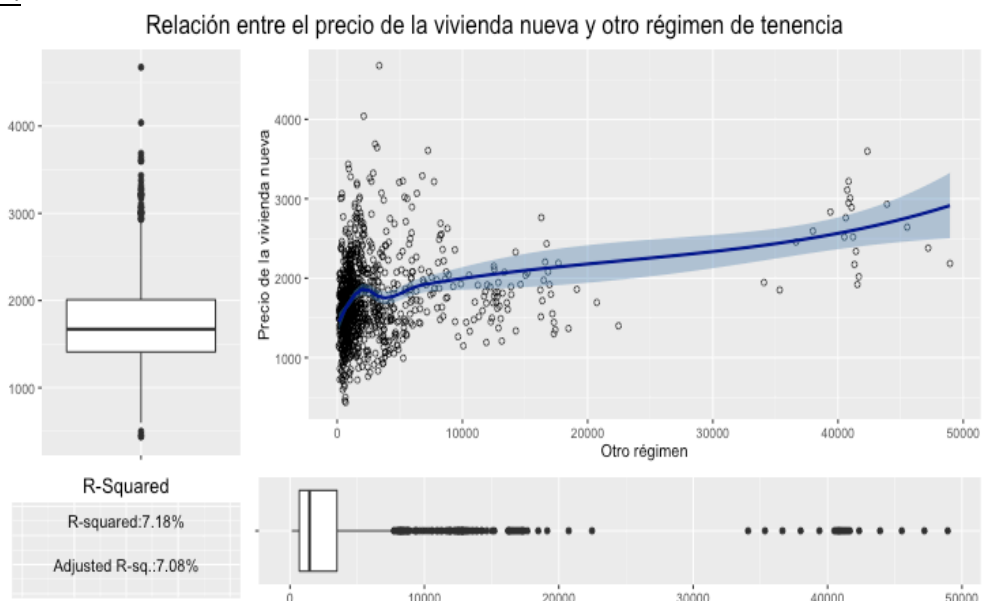
Fuente: Ministerio de Fomento. Elaboración propia.

En esta variable cuesta más de ver el efecto que tiene en relación al precio de la vivienda nueva, ya que la mayoría de áreas urbanas tiene un número de viviendas secundarias muy reducido y, dependiendo del sujeto de estudio, estas viviendas tienen un precio de vivienda nueva que puede tomar valores en un intervalo muy amplio.

En la relación entre ambas variables vemos que el número de viviendas secundarias sólo explica el 0.74% de la variancia sistemática de la variable respuesta, y la correlación lineal entre ambas variables es prácticamente nula y negativa, de valor -0.0858.

La cuarta variable presentada es el régimen de tenencia “otro régimen”:

**Figura 5.7:** Relación entre el precio de la vivienda nueva y el régimen de tenencia “otro régimen”.



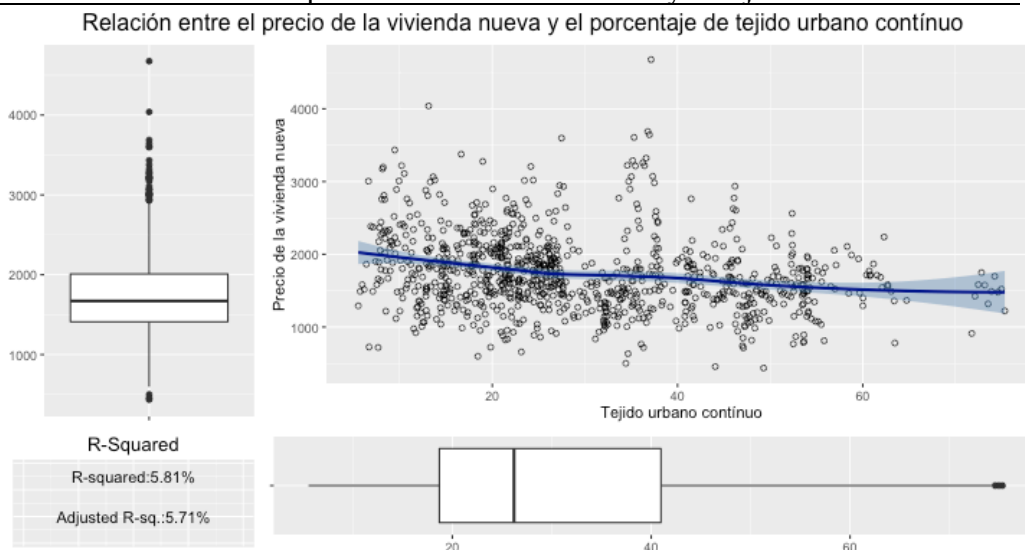
Fuente: Ministerio de Fomento. Elaboración propia.

Esta variable presenta un problema parecido al anterior, donde las áreas urbanas con mayor número de viviendas pueden tener muchas más viviendas de este tipo de régimen de tenencia.

La variable “otro régimen” de tenencia explica un 7.18% de la variancia sistemática de la variable “precio de la vivienda nueva”, y la correlación lineal entre ambas variables es de 0.2680.

La última variable presentada es el porcentaje de superficie declarada tejido urbano continuo:

**Figura 5.8:** Relación entre el precio de la vivienda nueva y el tejido urbano continuo.



Fuente: Ministerio de Fomento. Elaboración propia.

Con esta variable sí se observa gráficamente (ver Figura 5.8) un patrón en el comportamiento de una variable frente a la otra, ya que la variable que representa porcentualmente el tejido urbano continuo no depende del tamaño del parque de la vivienda de un área urbana como en las dos anteriores variables. Aun así, se observa una mayor presencia de *outliers* en aquellas observaciones de entre 0% y 40% de tejido urbano continuo.

La variable del “tejido urbano continuo” explica un 5.81% de la variancia sistemática del precio de la vivienda nueva, y la correlación lineal entre ambas variables es negativa, de -0.2411.

## 5.2. Análisis de los datos de panel

Siendo la variable de estudio el “Precio de la vivienda nueva”, y a partir de las variables obtenidas del Atlas Digital, se puede obtener un modelo con las variables que explican el comportamiento de la variable respuesta. El modelo presentado está elaborado con un total de N=946 observaciones, de n=86 sujetos a lo largo de un período de T=11 años.

**Tabla 5.1:** Modelo de panel para la explicación del precio de la vivienda nueva según área urbana.

<i>Variable</i>	$\beta_j$	$\chi^2$	<i>df</i>	<i>P-valor</i>
<i>Precio vivienda usada</i>	9.7750e-01	857.5124	1	<2.2e-16
<i>Tasación vivienda usada</i>	9.2336e-02	10.6474	1	0.0011022
<i>Viviendas secundarias</i>	-2.8449e-04	4.4200	1	0.0355206
<i>Otro régimen</i>	-2.4171e-02	5.8443	1	0.0156275
<i>Tejido urbano continuo</i>	-1.7962e+01	13.4700	1	0.0002424
$R^2$	0.59955			
$R^2_{adj}$	0.54187			

Fuente: Elaboración propia.

Junto con el modelo, se presentan también los efectos fijos de cada sujeto,  $\alpha_i$ , es decir, la aportación de cada sujeto de forma individual a la variable respuesta.

**Tabla 5.2:** Efecto fijo de cada área urbana sobre el modelo estimado.

<i>Área urbana</i>	$\alpha_i$	<i>Área urbana</i>	$\alpha_i$
<i>A Coruña</i>	571.4917	<i>Lleida</i>	944.3545
<i>Albacete</i>	365.1656	<i>Logroño</i>	558.8475
<i>Alcoy</i>	978.8694	<i>Lorca</i>	746.5831
<i>Alicante - Elche</i>	885.6139	<i>Lugo</i>	1208.3021
<i>Almería</i>	782.6300	<i>Madrid</i>	1143.7180
<i>Aranjuez</i>	221.0229	<i>Málaga</i>	833.2281
<i>Arrecife</i>	1130.4535	<i>Manresa</i>	1053.5895
<i>Asturias</i>	894.8289	<i>Melilla</i>	1081.3150
<i>Ávila</i>	385.5328	<i>Mérida</i>	951.7474
<i>Badajoz</i>	514.1260	<i>Motril</i>	907.0654
<i>Bahía de Algeciras</i>	791.7809	<i>Murcia</i>	848.4752
<i>Bahía de Cádiz</i>	806.0645	<i>Orihuela</i>	605.6365
<i>Barcelona</i>	1590.8852	<i>Ourense</i>	570.6664
<i>Bilbao</i>	909.1034	<i>Palencia</i>	420.3531
<i>Blanes – Lloret de Mar</i>	369.2546	<i>Palma</i>	839.5122
<i>Burgos</i>	277.7357	<i>Pamplona</i>	551.3532
<i>Cáceres</i>	568.4864	<i>Ponferrada</i>	709.6640
<i>Cartagena</i>	716.8024	<i>Sagunto/Sagunt</i>	654.4886
<i>Castellón de la Plana</i>	400.3253	<i>Salamanca</i>	554.1936
<i>Ceuta</i>	546.0849	<i>Sanlúcar de Barrameda</i>	986.5245
<i>Ciudad Real</i>	254.0518	<i>Sant Feliu de Guíxols</i>	345.4171
<i>Córdoba</i>	783.0610	<i>Santa Cruz de Tenerife – La Laguna</i>	1284.1149
<i>Costa Blanca</i>	482.3890	<i>Santander - Torrelavega</i>	436.5935
<i>Costa del Sol</i>	329.2561	<i>Santiago de Compostela</i>	832.6957
<i>Cuenca</i>	417.7639	<i>Segovia</i>	414.4269
<i>Denia – Jávea</i>	402.2015	<i>Sevilla</i>	1142.3690
<i>Donostia / San Sebastián</i>	754.5453	<i>Soria</i>	595.5543
<i>Eivissa</i>	682.9603	<i>Talavera de la Reina</i>	734.3603
<i>El Ejido</i>	798.6138	<i>Tarragona - Reus</i>	505.6911
<i>Elda – Petrer</i>	725.0392	<i>Tenerife Sur</i>	874.2557
<i>Ferrol</i>	426.5247	<i>Teruel</i>	226.9826
<i>Gandía</i>	1394.3733	<i>Toledo</i>	441.6460
<i>Girona</i>	1141.5421	<i>Torre Vieja</i>	340.5724
<i>Gran Canaria Sur</i>	907.3277	<i>Utrera</i>	1010.3578
<i>Granada</i>	983.6149	<i>Valencia</i>	1392.2987
<i>Guadalajara</i>	547.4717	<i>Valladolid</i>	283.4751
<i>Huelva</i>	621.2982	<i>Valle de la Orotava</i>	1305.7722
<i>Huesca</i>	344.9626	<i>Vélez – Málaga</i>	820.2926
<i>Jaén</i>	620.2151	<i>Vigo - Pontevedra</i>	829.6127
<i>Las Palmas de Gran Canaria</i>	1244.1467	<i>Vitoria / Gasteiz</i>	240.5994
<i>León</i>	332.9882	<i>Zamora</i>	977.1259
<i>Linares</i>	927.6034	<i>Zaragoza</i>	267.5462

Fuente: Elaboración propia.

Dado que las variables del modelo son todas continuas, la interpretación es la misma para cada una de ellas. Es decir, para cada unidad de una variable, el precio esperado medio de la vivienda nueva aumenta en los mismos puntos que la estimación del coeficiente poblacional asociado. Por ejemplo, dada un área urbana con un precio de la vivienda usada de 2000 euros, la contribución al precio esperado medio de la vivienda nueva de esta variable será de 1995.5 unidades.

Tomando como ejemplo el área urbana de Barcelona en el año 2014:

**Tabla 5.3:** Ejemplo de la estimación del precio de la vivienda nueva con el área urbana de Barcelona.

<i>Variable</i>	$\beta_j$	<i>Valor real medio</i>	<i>Contribución</i>
<i>Precio vivienda usada</i>	9.7750e-01	1591.1	1555.3003
<i>Tasación vivienda usada</i>	9.2336e-02	1644.5	153.6933
<i>Viviendas secundarias</i>	-2.8449e-04	1182167	-336.3147
<i>Otro régimen</i>	-2.4171e-02	34085.04	-823.8695
<i>Tejido urbano continuo</i>	-1.7962e+01	26.25625	-471.6148
<i>Total</i>			131.1946
$\alpha_i$	1590.8852		

Fuente: Elaboración propia.

El valor teórico del precio de la vivienda nueva es de 1772.08 euros (resultado de 1590.8852 + 131.1946), siendo el real observado 1945.6.

Tomando como segundo ejemplo el área urbana de Tarragona – Reus en el año 2014:

**Tabla 5.4:** Ejemplo de la estimación del precio de la vivienda nueva con el área urbana de Tarragona.

<i>Variable</i>	$\beta_j$	<i>Valor real</i>	<i>Contribución</i>
<i>Precio vivienda usada</i>	9.7750e-01	1138.5	1112.8838
<i>Tasación vivienda usada</i>	9.2336e-02	1146.7	105.8817
<i>Viviendas secundarias</i>	-2.8449e-04	464256.8	-132.0764
<i>Otro régimen</i>	-2.4171e-02	2407.516	-58.1921
<i>Tejido urbano continuo</i>	-1.7962e+01	12.838	-230.5962
<i>Total</i>			797.855
$\alpha_i$	505.6911		

Fuente: Elaboración propia.

El valor teórico del precio de la vivienda nueva es de 1303.55, siendo el real observado de 1423.5 euros.

### 5.3. Validación del modelo

#### 5.3.1. Test de heterogeneidad

El primer test realizado para validar el modelo es el de heterogeneidad presentado anteriormente que sirve para validar si el efecto de los efectos fijos de cada sujeto es distinto, y por lo tanto, el modelo estimado es bueno.

Con el estadístico F, de valor 256.016 con 5 y 855 grados de libertad, nos lleva a rechazar la hipótesis nula de no heterogeneidad con un p-valor asociado menor a  $2.22e-16$ , por lo tanto los coeficientes  $\alpha_i$  son validos.

#### 5.3.2. Supuesto de homoscedasticidad

Para validar este supuesto, se lleva a cabo el test de Breusch-Pagan, donde la hipótesis nula es la presencia de homoscedasticidad.

El estadístico del test toma valor 10.437 con 5 grados de libertad, con lo que nos lleva a no rechazar esta hipótesis y se confirma el supuesto de homoscedasticidad.

#### 5.3.3. Efectos fijos vs efectos aleatorios

Para contrastar ambos estimadores, se utiliza el test de Hausman a través de la comparación del modelo estimado (por efectos fijos) y del mismo modelo estimado por efectos aleatorios.

El estadístico del test toma valor 29.931 con 5 grados de libertad, lo que nos lleva a rechazar la hipótesis nula y a elegir el estimador de efectos fijos como mejor estimador.

## 6. SOLUCIONES Y CONCLUSIONES

Del tratamiento de las variables a través de los índices, se ha observado que el número de viviendas vacías, aun cuando hay mucha gente que no puede acceder a una, ha aumentado en algunas áreas urbanas mucho más que en otras. Sin embargo, al tratar el número de viviendas vacías per cápita, esta dispersión demuestra que el número de viviendas vacías per cápita es similar entre áreas urbanas, por lo que se puede concluir que el número de viviendas vacías ha aumentado más en aquellas áreas urbanas con mayor población, las de las grandes ciudades (Madrid, Barcelona, Valencia, Sevilla, Bilbao,...).

Al tratar otras variables per cápita, destaca la dispersión en las viviendas secundarias per cápita entre áreas urbanas. Esto es debido a áreas urbanas como la de Sant Feliu de Guíxols o Torrevejea, donde el número de viviendas secundarias es mucho mayor al número de viviendas principales. Esta característica se da en otras áreas urbanas con una misma característica: son de zona de costa (zonas de vacaciones, turismo).

Otra conclusión importante a la que se llega con los datos, para luego analizar posibles soluciones a los problemas del mercado de la vivienda español, es que las viviendas con régimen de tenencia de propiedad han aumentado en la mayoría de áreas urbanas entre 2004 y 2014, mientras que las viviendas en alquiler han disminuido en todas las áreas urbanas, aunque estas últimas han crecido lentamente los tres últimos años.

Finalmente, y según el modelo estimado, se puede concluir que el precio de la vivienda nueva depende, en gran parte, del precio de la vivienda ya usada y de la tasación de esta.

El valor de la vivienda, como se planteaba al principio de este trabajo, es el que determina el uso de esta. Según la rentabilidad ofrecida, la vivienda será destinada al mercado de la propiedad, del alquiler o de la vivienda protegida. La mayor rentabilidad se obtiene a través de la venta de una vivienda, por eso la mayoría de viviendas están destinadas al mercado de la propiedad, con el claro aumento de viviendas para uso secundario y el número de viviendas vacías. Por lo tanto, se puede concluir que el mercado de la propiedad está destinado a personas con rentas medias y altas, las cuales pueden permitirse pagar una vivienda en propiedad e incluso una vivienda secundaria, mientras que las personas con rentas bajas quedan excluidas de este mercado, y deben recurrir a otras formas de acceso a la vivienda, como pueden ser el alquiler o la vivienda de protección oficial. Con esta situación definida, se observa una sobreoferta en el mercado de la propiedad, mientras que los mercados del alquiler y la vivienda de protección oficial, aunque en los diez últimos años hayan disminuido su oferta, sí se observa cómo la demanda en los últimos años de crisis ha ido aumentando. Por lo tanto, el mercado de la vivienda español se define como el conjunto de tres mercados en desequilibrio donde en uno, el de la propiedad, la oferta es muy superior a la demanda, y los otros dos, los del alquiler y la vivienda de protección oficial,



tienen una demanda por encima de la oferta. Esta situación lleva a concluir que existe una necesidad de trasladar viviendas del mercado de la propiedad a los mercados del alquiler y de la vivienda de protección oficial, sobretodo en las grandes ciudades, donde el número de viviendas vacías es superior al resto de áreas urbanas.

El problema que se plantea aquí es como trasladar las viviendas de un mercado al otro. Lo más importante, en este problema, es el garantizar el acceso de toda la población a una vivienda. Por lo tanto, se debe priorizar la visión redistributiva del problema, sin dejar a un lado la búsqueda de la visión asignativa del problema, con lo que debe crearse un mercado del alquiler fuerte, con precios competitivos, a los que pueda acceder la mayor parte de la población y, finalmente, garantizar la vivienda para aquellos que no puedan pagar un alquiler a través de la vivienda de protección oficial. Sin embargo, el precio del alquiler depende del precio de la vivienda, el cual depende de los precios del mercado de la propiedad, los cuales han crecido mucho debido a la burbuja inmobiliaria de 1997 a 2007. Por eso, el precio del alquiler es elevado, pero los gastos que supone no dan la misma rentabilidad al propietario de la vivienda que la opción de venderla, y por eso en muchas ocasiones la vivienda queda vacía antes que puesta en alquiler.

A corto y medio plazo una de las medidas que se pueden aplicar para que la vivienda sea puesta en alquiler son medidas relacionadas con la fiscalidad, como desgravaciones en el IRPF o reducciones de impuestos locales. Estas medidas deberían conseguir que la rentabilidad obtenida a través de la renta del alquiler fuera similar a la rentabilidad obtenida con la venta del inmueble. Dado que algunas medidas de estas ya existen en la actualidad, se puede desincentivar la compra de vivienda a través de retirar ayudas existentes sobre estas, provocando una caída en la rentabilidad de estas operaciones.

Pero lo importante es el dar acceso a la vivienda a personas que no pueden permitírselo. Para ello, también se puede recurrir a medidas fiscales, como la desgravación en el IRPF de parte de la renta del alquiler.

Medidas como estas deberían provocar que el precio del alquiler fuera representativamente inferior al precio de la compra de una vivienda, y la oferta debería aumentar, garantizando la vivienda a un colectivo mayor. Sin embargo, para aquellas personas que no puedan pagar un alquiler, el sector público debería incentivar y controlar la creación de un parque de vivienda protegida oficial, destinada a un *target* concreto, aquellas personas que no pueden permitirse la vivienda a través de la propiedad o el alquiler.

Al igual que en el mercado del alquiler, para favorecer el acceso a la vivienda de protección oficial la primera actuación que debe llevarse a cabo es el aumento del parque de vivienda social, preferiblemente a través de la transformación de la vivienda que se encuentra en el mercado en vivienda de protección oficial, así se conseguiría reducir la sobreoferta del

mercado y reducir el número de viviendas vacías. Para conseguir ese traspaso de un mercado al otro, el sector público puede recurrir a los incentivos fiscales, junto con las ayudas directas. Estas ayudas directas, además, repercuten en aquellos que vivirán en la vivienda de protección oficial, ya que el precio que pagarán por la vivienda es inferior al precio de mercado.

## BIBLIOGRAFÍA

1. ACHIM ZEILEIS. 2014. *ineq: Measuring Inequality, Concentration, and Poverty. R Package version 0.2-13*. <https://CRAN.R-project.org/package=ineq>
2. BERGES, ÁNGEL; ONTIVEROS, EMILIO. 2007. La nueva Ley de Suelo desde la perspectiva económica. Ministerio de Fomento.
3. BORGIA SORROSAL, SOFÍA; DELGADO GIL, ANDRÉS. 2009. Evolución de las políticas de vivienda en España. Comparativa con la UE-15. Instituto de Estudios Fiscales.
4. BREUSCH, T. S.; PAGAN, A.R. 1979. A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 47, 1287-1294.
5. CASTAÑO MARTINEZ, MARÍA SOLEDAD. 2012. La eficacia de la política de la vivienda en España. *Revista ICE*.
6. COWELL, FRANK A. 2009. *Measuring Inequality*. Oxford University Press.
7. FERNÁNDEZ MORALES, ANTONIO. 1991. Estimación de índices de desigualdad y pobreza a través del ajuste económico de curvas de Lorenz. Universidad de Málaga.
8. FREES, EDWARD W. 2003. *Longitudinal and Panel Data: Analysis and Applications for the Social Sciences*. Cambridge University Press.
9. GARCÍA MONTALVO, JOSÉ. 1999. El precio del suelo: la polémica interminable. IVIE y Universidad de Valencia.
10. GOERLICH, FRANCISCO J. 1998. Desigualdad, diversidad y convergencia: (algunos) instrumentos de medida. Instituto Valenciano de Investigaciones Económicas.
11. Wickham. 2009. *ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. *R package*. <https://cran.r-project.org/web/packages/ggplot2/index.html>
12. KOLM, SERGE-CHRISTOPHE. 1975. Unequal Inequalities. I & II. School for Higher Studies in the Social Sciences and CEPREMAP.
13. LEAL, JESÚS. 2005. La política de vivienda en España. Cáritas.
14. Ley 29/94, de 24-11-1994 (BOE de 25/11/1994).
15. MANCERO, XAVIER. 2000. Revisión de algunos indicadores para medir la desigualdad. CEPAL.
16. MONTERO, R. 2005. Test de Hausman. Documentos de Trabajo en Economía Aplicada. Universidad de Granada.
17. OBSERVATORIO DE VIVIENDA Y SUELO. 2013. Boletín Especial de Rehabilitación y Alquiler. Centro de Publicaciones, Ministerio de Fomento.
18. PAREJA EASTAWAY, MONTSERRAT; SÁNCHEZ MARTINEZ, MARÍA TERESA. 2012. La Política de vivienda en España: lecciones aprendidas y retos de futuro. *Revista Galega de Economía*.
19. PITTINI, ALICE; LAINO, ELSA. 2011. Housing Europe Review 2012. The nuts and bolt of European social housing systems. CECODHAS Housing Europe's Observatory.

20. Plan estatal de vivienda y rehabilitación 2009-2012. Guía práctica. 2008. Ministerio de vivienda.
21. Real Decreto 1346/1976, de 09-04-1976 (BOE de 17/06/1976).
22. Real Decreto 2066/2008, de 12-12-2008 (BOE de 24/12/2008).
23. Real Decreto Legislativo 7/2015 de 30-10-2015 (BOE de 31/10/1976).
24. RODRÍGUEZ, RAQUEL. 2010. La política de vivienda en España en el contexto europeo. Deudas y retos. Revista INVI, 25, 69,124-159.
25. RODRÍGUEZ LÓPEZ, JÚLIO. 2016. El mercado de vivienda en España. En línea: <http://julrodlop.blogspot.com.es/2016/01/el-mercado-de-vivienda-en-espana-en.html>
26. STEVENSON, SIMON; YOUNG, JAMES. 2006. Comportamiento especulativo en los mercados de vivienda: una perspectiva internacional. Papeles de Economía Española.
27. TORRES-REYNA, OSCAR. 2007. Panel Data Anlaysia Fixed and Random Effects using Stata (v 4.2). Universidad de Princeton.
28. VINUESA ANGULO, JULIO. 2012. Estimación de las necesidades de vivienda en España. Juvian01.

## ANEXO

### A.1 Definición de variables

**Tabla A.1:** Definición de las variables usadas.

<i>Tipo de variable</i>	<i>Variable</i>	<i>Definición</i>
<i>Demográfica</i>	Población	Número de personas por nivel geográfico.
	Densidad de población	Número de habitantes por kilómetro cuadrado por nivel geográfico.
	Extranjeros	Número de personas de nacionalidad extranjera por nivel geográfico.
	Edad media de la población	Media aritmética del conjunto de edades de los habitantes.
<i>Mercado de la vivienda</i>	Precio vivienda libre total	Precio de la vivienda libre, independientemente de su clasificación.
	Precio vivienda libre nueva	Precio de la vivienda libre nueva
	Precio vivienda libre usada	Precio de la vivienda libre usada
	Valor tasado vivienda total	Valor de tasación de la vivienda, independientemente de su clasificación.
	Valor tasado vivienda nueva	Valor de tasación de la vivienda nueva.
	Valor tasado vivienda usada	Valor de tasación de la vivienda usada.
	Número de transacciones anuales	Número de transacciones que se llevan a cabo en el mercado de la vivienda en un año concreto por nivel geográfico.
	Número de transacciones trimestrales	Número de transacciones que se llevan a cabo en el mercado de la vivienda en un trimestre concreto por nivel geográfico.
<i>Parque de la vivienda</i>	Densidad de viviendas familiares	Número de viviendas familiares por kilómetro.
	Densidad de viviendas familiares principales	Número medio de viviendas familiares principales por kilómetro.
	Densidad de viviendas secundarias	Número medio de viviendas secundarias por kilómetro.
	Densidad de viviendas vacías	Número medio de viviendas vacías por kilómetro.
	Viviendas familiares principales	Una vivienda es considerada familiar cuando es utilizada toda o la mayor parte del año como residencia habitual de una o más personas.
	Viviendas secundarias	Una vivienda familiar se considera secundaria cuando es utilizada solamente parte del año, de forma estacional, periódica o esporádica y no constituye residencia habitual de una o varias personas.
	Viviendas vacías	Una vivienda familiar se considera desocupada o vacía cuando, sin encontrarse en ninguna de las situaciones

	anteriormente consideradas, habitualmente se considera deshabitada.
Otras viviendas	Resto de viviendas que no se incluyen en las definiciones anteriores.
Viviendas familiares principales convencionales en régimen de tenencia de propiedad pagada	Una vivienda se considera en propiedad cuando es propietaria alguna de las personas que la ocupa. Por extensión, cuando el titular ha fallecido, emigrado, etc. También se pueden considerar propietarios aquellos que, no siendo aún titulares legales de la vivienda, lo son a modo natural por vía de transmisión, herencia u otra modalidad.
Viviendas familiares principales convencionales en régimen de tenencia de propiedad con pagos pendiente	Se considera que la vivienda tiene pagos pendientes cuando existe un compromiso, escrito o verbal, de efectuar pagos aplazados por la adquisición de la misma (por ejemplo, préstamos hipotecarios o de cualquier otra índole ligados a la adquisición de la vivienda).
Viviendas familiares principales convencionales en régimen de tenencia de propiedad por herencia o donación	Se contempla esta forma de acceso a la propiedad con independencia de que la vivienda pueda tener sobre sí cualquier tipo de gravamen.
Viviendas familiares principales convencionales en régimen de tenencia de alquiler	Una vivienda se considera en alquiler cuando, estando ocupada por los miembros de un hogar, alguno de sus componentes satisface por su uso una cantidad anual o mensual, en metálico o en especie, e independientemente de que medie contrato o no. El hecho de que esté alquilado con o sin muebles deberá constar en el contrato o declararse por algún otro medio.
Viviendas familiares principales convencionales en régimen de tenencia gratuito	Una vivienda es facilitada gratuita o semigratuitamente por el patrono, organismo o empresa donde trabaja alguno de los miembros del hogar, cuando lo es en virtud de contrato de empleo, aún cuando dicho miembro haya fallecido, esté jubilado o no trabaje actualmente para dicho patrono, organismo o empresa, siempre y cuando la vivienda siga perteneciendo al mismo. También se incluyen en esta rúbrica las viviendas familiares cuyo uso es facilitado como remuneración en especie. Una vivienda es facilitada gratuitamente o semigratuitamente por otras personas o instituciones cuando, sin ser la vivienda propiedad de ninguno de los ocupantes ni estar en régimen de alquiler, no existe contraprestación alguna por el uso de la vivienda o bien la contraprestación está fijada, de forma voluntaria por el propietario de la vivienda, por debajo del precio de alquiler del mercado. No se consideran aquí, por tanto, los alquileres de renta antigua cuyo precio está, también, aunque no voluntariamente, por debajo de su valor.
Viviendas familiares principales convencionales en otro régimen de tenencia	Una vivienda se tiene en otra forma si su régimen de tenencia no se ajusta a ninguna de las modalidades siguientes: propiedad, alquiler o cedida gratis.
Ocupación del Superficie de suelo artificial	Hectáreas de suelo artificial por nivel geográfico.

<i>suelo</i>		
	Tejido urbano continuo	Proporción de tejido urbano continuo sobre el total de suelo artificial
	Zonas industriales y comerciales	Proporción de zonas industriales y comerciales sobre el total de suelo artificial.
	Redes viarias y ferroviarias	Proporción de redes viarias y ferroviarias sobre el total de suelo artificial.
	Zonas portuarias	Proporción de zonas portuarias sobre el total de suelo artificial.
	Aeropuertos	Proporción de suelo dedicado a aeropuertos sobre el total de suelo artificial.
	Zona de extracción minera	Proporción de zonas de extracción minera sobre el total de suelo artificial.
	Escombreras y vertederos	Proporción de suelo dedicado a escombreras y vertederos sobre el total de suelo artificial.
	Zonas en construcción	Proporción de zonas en construcción sobre el total de suelo artificial.
	Zonas verdes urbanas	Proporción de zonas verdes urbanas sobre el total de suelo artificial.
	Instalaciones deportivas y recreativas	Proporción de suelo dedicado a instalaciones deportivas y recreativas sobre el total de suelo artificial.

Fuente: Ministerio de Fomento. Elaboración propia.

## A.2. Ejemplos del código implementado en R

### A.2.1. Carga de datos y de los paquetes necesarios

```
load("~/Desktop/TFG/Dades/Data.RData")
library("xlsx")

library(car)

library(doBy)

library(gplots)

library(foreign)
library(graphics)
library("plm")

library(ggplot2)

library(gridExtra)

library(reshape)
library(ineq)
library(lme4)

library(lattice)
```

### A.2.2. Lectura de datos

Los datos se han descargado del Atlas Digital según tipo de variable (demográfica, del mercado de la vivienda, parque de la vivienda o ocupación del suelo), según región (comunidad autónoma, área urbana o municipio) y por año. En este ejemplo, se presenta la lectura del archivo del mercado de la vivienda de las áreas urbanas del año 2004, junto con la transformación de los valores *missing* (representados por 999999999 en el archivo Excel) en NA para tratarlos en R.

```
mercviv_au04<-read.xlsx("mercviv_areaurbana04.xls",sheetName=2)
names(mercviv_au04)<-c("id","AreaUrbana", "Prec.viv.nueva", "Prec.viv.us
ada","Trans.anuales")
mercviv_au04<-mercviv_au04[order(mercviv_au04$AreaUrbana),]
for(iin 1:length(mercviv_au04[,1])){
  for(j in 3:length(mercviv_au04[1,])){
    if(mercviv_au04[i,j]==999999999){
      mercviv_au04[i,j]<-NA
    }
  }
}
```

### A.2.3. Imputación de *missings*

Se imputa la media de cada variable a los valores *missing*. En el ejemplo se tratan los missing de los datos del mercado de la vivienda para las áreas urbanas en 2014.

```
med<-mercviv_au14
for(i in 1:length(mercviv_au14[,1])){
  for(j in 1:length(mercviv_au14[1,])){
    if(is.na(mercviv_au14[i,j])) mercviv_au14[i,j]<-median(med[,j][!i
s.na(med[,j])])
  }
}
```

### A.2.4. Interpolación de datos

Mediante el método de la interpolación, se obtienen los datos para los años que no están disponibles. En este ejemplo, se obtienen los datos del parque de la vivienda para las áreas urbanas en 2004 a partir de los datos del parque de la vivienda de las áreas urbanas en 2001 y 2011.

```
parcviv_au04<-parcviv_au11
for(i in 1:length(parcviv_au11[,1])){
  for(j in 3:length(parcviv_au11[3,])){
```



```

    parcviv_au04[i,j]<-parcviv_au01[i,j]*(parcviv_au11[i,j]/parcviv_
au01[i,j])^(3/10)
  }
}

```

#### A.2.5. Creación de las bases de datos anuales

Se juntan, mediante la función *merge*, los datos anuales de las variables demográficas, del mercado de la vivienda, del parque de la vivienda y de la ocupación del suelo. En el ejemplo posterior, se muestra la creación de la base de datos de 2014. En esta, se le añade una columna con el año al cual pertenecen los datos.

```

####Datos 2014
data14<-merge(pob_au14[, -7],mercviv_au14[,c(1:7)], by=c("id", "AreaUrba
na"))
data14<-merge(data14,parcviv_au14, by=c("id", "AreaUrbana"))
data14<-merge(data14,ocu_au14,by=c("id", "AreaUrbana"))
summary(data14)
data14<-cbind(data14,rep(2014,86))
names(data14)<-c("id", "AreaUrbana", "Poblacion", "Dens.Pob", "Extranjeros", "
EdadMedia", "Prec.viv.nueva", "Prec.viv.usada", "Tas.viv.nueva", "Tas.viv.usa
da", "Trans.anuales", "Dens.viv.fam", "Dens.vivi.fam.princ", "Dens.viv.sec", "D
ens.viv.vacias", "Viv.fam.princ", "Viv.sec", "Viv.vaci", "Otras.viv", "Prop.pag
ada", "Pagos.pend", "Prop.por.herdo", "Alquiler", "Gratis", "Otro.regimen", "S
up.Suelo.Art", "Tej.Urb.Cont", "Tej.Urb.Disc", "Aeropuerto", "ZonaPort", "Año")

```

#### A.2.6. Creación de la base de datos de panel

Se juntan las bases de datos de cada año para crear la base de datos de panel para el período entre 2004 y 2014.

```

data<-NULL
data<-rbind(data04,data05,data06,data07,data08,data09,data10,data11,
data12,data13,data14)
summary(data)

```

#### A.2.7. Ejemplos de gráficos del trabajo

- Gráfico de la relación entre el precio de la vivienda nueva y la tasación de la vivienda usada, con el coeficiente  $R^2$  (ver Figura 4.5).

```

Scat <- ggplot(data = data, aes(x=Tas.viv.usada, y=Prec.viv.nueva)) +
geom_point(shape=1, color = 'black', alpha=0.8) + xlab('Tasación de la vi
vienda usada') + ylab('Precio de la vivienda nueva') + geom_smooth(color
= 'darkblue', fill='steelblue', size = 1, method=loess, se = TRUE, formu

```

```

la = y~x, fullrange=TRUE) +theme(axis.title = element_text(size=11))
BP1 <- ggplot(data = data, aes(x=factor(1), y=Prec.viv.nueva)) +geom_b
oxplot() + xlab(NULL) + ylab(NULL) + theme(axis.text.x=element_blank())
+ theme(axis.title = element_text(size=9))
BP2 <- ggplot(data = data, aes(x=factor(1), y=Tas.viv.usada)) +theme(a
xis.title = element_text(size=9)) +geom_boxplot() + ylab(NULL) + xlab(NUL
L) + theme(axis.text.y=element_blank()) +coord_flip()

lm <- summary(lm(Prec.viv.nueva~Tas.viv.usada, data))
lm$r.squared
lm$adj.r.squared

R <- ggplot() + ggtitle('R-Squared') +
  annotate("text", x=1,y=1.5, label= "R-squared:77.23%") +
  annotate("text", x=1, y=1.2, label= "Adjusted R-sq.:76.26%") +
  ylim(1,1.6) + xlim(0.8,1.2) + theme(axis.text=element_blank(), axis.lin
e = element_blank(),axis.ticks = element_blank()) +ylab(NULL) +xlab(NULL)

grid.arrange(grobs = list(BP1, Scat, BP2, R),
  layout_matrix = MAT,top = "Relación entre el precio de la v
ivienda nueva y la tasación de la vivienda usada")

```

- Gráfico de la evolución del precio del suelo (ver Figura 1.11):

```

#Evolución precio suelo
colnames(preciosuelo) <- c('Año', 'Precio')
ggplot(preciosuelo, aes(x=Año, y=Precio)) +
  geom_point(shape=19, color='steelblue') +
  geom_line(color='steelblue') +
  ylab(expression('Euros por m'^2)) +
  xlab('Año') +
  ggtitle('Evolución del precio del suelo en España') +
  theme(plot.title = element_text(size=13)) +
  theme(axis.title = element_text(size=11))

```

- Gráfico del número de viviendas vacías (ver Figura 2.29).

```

ggplot(parcviv_ccaa, aes(x=CCAA, y = Viv.vaci)) +
  geom_bar(stat = 'identity') + ylab('Número de viviendas') +
  theme(axis.text.x = element_text(angle = 90, hjust = 1)) +
  ggtitle('Número de viviendas vacías por CCAA') +

```

```
theme(plot.title = element_text(size=13)) +
theme(axis.title = element_text(size=11))
```

#### A.2.8. Estimación del modelo con datos de panel

Estimación del modelo de regresión con datos de panel usando el paquete “plm” de R. Junto a la estimación, se presentan los test de Breusch-Pagan y de Hausman.

```
panel<-plm(Prec.viv.nueva~Prec.viv.usada+Tas.viv.usada+Viv.sec+Otro.
regimen+Tej.Urb.Cont,data=data, index=c("AreaUrbana","Año"),model="with
in")
summary(panel)
Anova(panel)
##Efectos fijos
ai<-fixef(panel)
##Diagnósticos del modelo
panel2<-plm(Prec.viv.nueva~Prec.viv.usada+Tas.viv.usada+Viv.sec+Otro.
.regimen+Tej.Urb.Cont,data=data, index=c("AreaUrbana","Año"),model="ran
dom")
summary(panel2)
## Test de Breusch-Pagan
bptest(panel,data=data)
## Hausman test
phtest(panel, panel2)
```

#### A.2.9. Cálculo de la desigualdad

Se presenta la función para el cálculo de las medidas de desigualdad a partir de las funciones obtenidas del paquete de R “ineq”. Se pone como ejemplo el cálculo de las medidas de desigualdad para el año 2014, donde en “descriptivos14” se encuentran el rango, la variancia y el coeficiente de variación y en índices14 los índices de Gini, Atkinson, Theil normalizado, Kolm y la entropía generalizada.

```
desigualdad<-function(x){
  des<-c((max(x)-min(x)),var(x),var.coeff(x),Gini(x),Atkinson(x),Theil(
x)/log(86),Kolm(x),entropy(x))
  return(des)
}

indices14<-NULL
for(i in 3:length(data14[1,])){
  indices14<-rbind(indices14,desigualdad(data14[,i]))
}
```

```
indices14<-indices14[-29,]
indices14<-data.frame(indices14,row.names=c("Poblacion","Dens.Pob","Extr
anjeros","EdadMedia","Prec.viv.nueva","Prec.viv.usada","Tas.viv.nueva","
Tas.viv.usada","Trans.anuales","Trans.trim","Dif.PreTas.nueva","Dif.PreTas
.usada","Dens.viv.fam","Dens.vivi.fam.princ","Dens.viv.sec","Dens.viv.vacia
s","Viv.fam.princ","Viv.sec","Viv.vaci","Otras.viv","Prop.pagada","Pagos.p
end","Prop.por.herdon","Alquiler","Gratis","Otro.regimen","Aeropuerto","
ZonaPort"))
colnames(indices14)<-c("Rango","Variancia","Coef. Var","Gini","Atkinson"
,"Theil","Kolm","Entropia Generalizada")
descriptivos14<-indices14[,1:3]
indices14<-indices14[,4:8]
```