



# **Análisis de la estructura espacial del precio de la vivienda en la ciudad de Madrid**

**Daniel Robles Quiñonero**  
Director: Fernando A. López Hernández

Trabajo Fin de Máster para la obtención del título de Máster en  
Administración y Dirección de Empresas (MBA)  
Curso 2015/2016

## **Indice**

RESUMEN.....	3
SUMMARY.....	4
INTRODUCCIÓN.....	5
El mercado inmobiliario.....	5
El mercado de la vivienda en España y su evolución temporal.....	5
Evolución y análisis del precio de la vivienda en la ciudad de Madrid.....	8
Los ciclos del mercado de la vivienda.....	10
El valor catastral y el precio de mercado.....	13
Factores determinantes del precio de la vivienda.....	15
JUSTIFICACIÓN DEL TRABAJO.....	16
Objetivos y propósito de la investigación.....	16
ESTADO DE LA CUESTIÓN.....	17
METODOLOGÍA.....	22
La econometría espacial.....	22
La georreferenciación y los sistemas de información geográfica (SIG).....	24
Las regresiones hedónicas y espaciales.....	33
Autocorrelación espacial global y local.....	37
DATOS.....	40
Fuentes de datos sobre precios de la vivienda.....	40
Base de datos utilizada en la investigación.....	45
Fuentes de datos sobre elementos georreferenciados.....	49
RESULTADOS.....	51
Análisis descriptivo de los datos.....	51
Análisis exploratorio de datos espaciales (AEDE).....	56
Autocorrelación espacial global (índice de Moran).....	57
Autocorrelación espacial local (Local Indicator of Spatial Association -LISA-).....	58
Resultados de las regresiones y análisis de coeficientes.....	60
Resultados y análisis de los mapas de <i>ARCGIS</i> .....	67
CONCLUSIONES.....	74
INDICE DE GRÁFICOS, TABLAS E IMÁGENES.....	79
BIBLIOGRAFÍA.....	82
Anexo 1: precio de mercado y valor catastral por barrios.....	85
Anexo 2: elementos georreferenciados con el SIG <i>ArcGis online</i> .....	90

## RESUMEN

El sector inmobiliario y de la construcción de vivienda tiene un peso indudable en la economía de cualquier país. Como sector puede representar porcentajes del PIB que varían entre el 5% y el 12% y es un sector con fuerte capacidad de arrastre en términos de empleo generado directa e indirectamente.

La vivienda es, sin duda, una de las inversiones más importantes que se realizan en la vida y el valor de transacción de la misma está sujeto a ciclos económicos y a múltiples variables asociadas tanto a la ubicación como a las características propias del inmueble. El precio de la vivienda es, por tanto, uno de los aspectos de mayor repercusión social en las economías desarrolladas, y por ello es objeto de estudio y análisis desde distintos puntos de vista.

Desde el enfoque econométrico, el análisis del precio de la vivienda se aborda desde el estudio de numerosas y muy diversas variables, elaborando modelos que permitan conocer el valor de transacción de bienes inmuebles. Mediante técnicas estadísticas se pretende conocer qué factores influyen en la formación de dichos precios y en qué medida. Para ello, en las investigaciones en este campo se tienen en cuenta tanto factores *indoor* propios del inmueble (tamaño o número de estancias) como factores externos *outdoor*, que permiten abordar la formación de los precios teniendo en cuenta un enfoque espacial, considerando atributos asociados a la localización de los inmuebles y a la proximidad de éstos a lugares o puntos de interés (parques, colegios, hospitales, nudos de comunicación, etc).

En este trabajo, haciendo uso de la información de un conocido portal inmobiliario y de herramientas informáticas y estadísticas específicas como los sistemas de información geográfica (SIG) se pretende elaborar un modelo que incorpore dichos factores espaciales. El objetivo principal de la investigación consiste en analizar la estructura espacial del precio de la vivienda en la ciudad de Madrid, así como las diferencias entre distritos y barrios y la influencia de determinadas variables de carácter espacial en la formación de los precios.

Partiendo de la premisa de la primera ley de la geografía elaborada por el geógrafo Waldo Tobler en 1970, que enunciaba *“todas las cosas están relacionadas con todo lo demás, pero las cosas cercanas entre sí lo están más que las distantes”* se pretende con este trabajo indagar y elaborar un modelo econométrico capaz de predecir el precio de las viviendas teniendo en cuenta tanto los factores intrínsecos de la propia vivienda como las variables espaciales.

**Palabras clave:** precio de la vivienda, mercado inmobiliario, geoestadística, econometría espacial.

## SUMMARY

The real estate and housing construction sector has an undeniable weight in the economy of any country. As a sector may represent a percentage of GDP ranging between 5% and 12%. Besides, it is a sector with strong pulling capacity in terms of direct and indirect employment generated.

Housing is undoubtedly one of the most important investments made in life and the transaction value of the same is subject to economic cycles and multiple variables associated with both the location and the characteristics of the property. The price of housing is therefore one of the aspects of greater social impact in developed economies, and therefore is a field of study and analysis from different points of view.

From the econometric point of view, the analysis of housing prices is approached from the study of numerous and diverse variables, developing econometric models which show the value of real estate transaction. By using statistical techniques it is aimed to know what factors influence the formation of such rates and to what extent. To this end, research in this field take into account both *indoor* factors (size or number of rooms) and *outdoor* factors, which can address the formation of prices taking into account a spatial approach, considering attributes associated with the location of the property and the proximity to any point of interest (parks, schools, hospitals, communication hubs, etc.).

In this research work, using information from a well-known real estate website and specific statistics tools as well as geographic information systems (GIS) it aims to develop a model that incorporates spatial factors. The main objective of the research is to analyze the spatial structure of housing prices in the city of Madrid, as well as differences between districts and neighborhoods and the influence of certain spatial variables in price formation.

Based on the premise of the first law of geography developed by the geographer Waldo Tobler in 1970, who enunciated "*everything is related to everything else, but near things are more related than distant things*" the aim with this work is to develop an econometric model capable of predicting the price of housing taking into account both the indoor factors of the property itself and the outdoor spatial factors.

**Keywords:** housing price, real estate market, geostatistics, spatial econometrics.

## INTRODUCCIÓN

### El mercado inmobiliario

El diccionario de la Real Academia Española define el término *inmobiliario/a* como la empresa o sociedad que se dedica a construir, arrendar, vender y administrar viviendas, pero el mercado o sector inmobiliario es aquél que comprende y desarrolla las fases del proceso productivo inmobiliario, es decir, el sector económico que opera en el ámbito de los bienes inmuebles, entendiéndose como tales las propiedades que no pueden moverse del lugar en el que están, tales como tierras, locales o viviendas.

En el ámbito de cualquier sociedad, el sector inmobiliario es uno de los mercados de mayor peso y relevancia, tanto por el peso relativo que representa sobre el cómputo global de una economía desarrollada como por otros factores directos o indirectos de igual importancia, como pueden ser el número de empleos que genera o la capacidad de “arrastre” que posee al involucrar a multitud de empresas y de sectores productivos en el seno de su actividad.

Es común en el estudio de la historia de las civilizaciones analizar la cantidad de inversiones realizadas en lo que a construcciones e infraestructuras se refiere, ya que por regla general las sociedades prósperas suelen invertir en construcción en todas sus ramas y ámbitos, desde la edificación residencial hasta las grandes infraestructuras civiles (puentes, acueductos, vías ferroviarias, monumentos, etc.) De este modo, desde las pirámides de Egipto al poderoso ímpetu constructor del Imperio romano, las sociedades reflejan sus períodos de bonanza económica traduciéndolo en construcciones de todo tipo, y España es un claro ejemplo de ello.

### El mercado de la vivienda en España y su evolución temporal

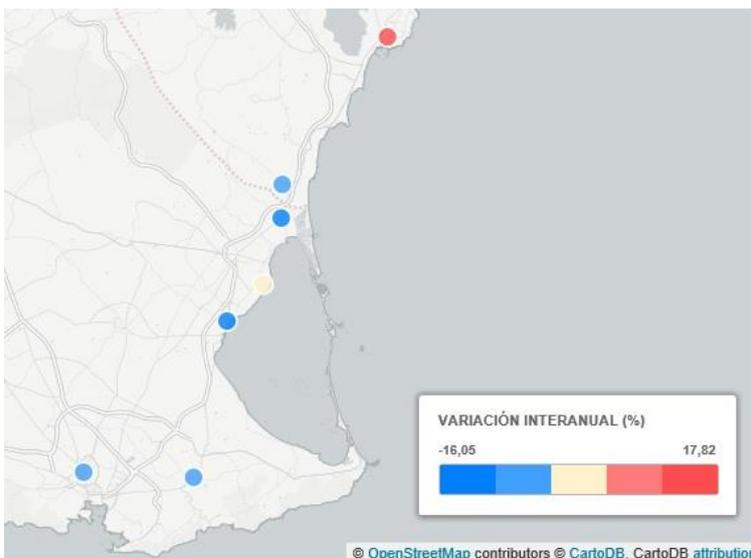
El mercado de la vivienda en España, y el sector inmobiliario y de la construcción en general, ha representado un peso significativamente alto en la economía del país. Si bien todas las sociedades dedican buena parte de sus recursos a la construcción en todas sus vertientes, el caso español es bastante definitorio. Por centrarnos en un período reciente, desde los años 50-60 del siglo XX España ha sufrido un importantísimo cambio en cuanto a políticas de vivienda y urbanismo, lo cual se ha cristalizado en profundos cambios tanto en ciudades como en la costa con el consiguiente cambio paisajístico e impacto medioambiental. Desde la construcción de barrios y ensanches en los años 70 y 80, en toda España se ha experimentado una migración constante del campo a la ciudad, transformando por completo la estructura de las urbes y haciendo incrementar la población de las mismas de una manera notable.

Junto con los cambios demográficos y las tendencias de migración hacia los ámbitos urbanos, el precio de la vivienda en España ha ido experimentando cambios a lo largo de todo este período, con épocas de ligeras subidas alternadas con otras de leves bajadas, pero todo ello de una manera más o menos estable. Todo ello comenzó a cambiar a finales de la década

de los 90 con políticas de suelo (cabe recordar que las políticas sobre vivienda siempre van acompañadas inevitablemente de políticas de suelo y de ordenación urbana) que condujeron a una liberalización de grandes cantidades de suelo acompañadas de grandes procesos de recalificación de suelos. Esto condujo, junto con el período de crecimiento económico que fue de 1996 a 2007, a lo que se ha venido a denominar después como el período de la burbuja inmobiliaria, unos años donde el precio de la vivienda comenzó a subir estrepitosamente y donde una serie de condiciones se dieron para favorecer una puesta en marcha de todo un sector (el de la construcción) que llegó a crecer de tal manera que acabó representando en el año 2007 casi el 13% del PIB español, cuando lo normal en un país de las características de España es que se mantenga rondando valores del 6 al 8%.

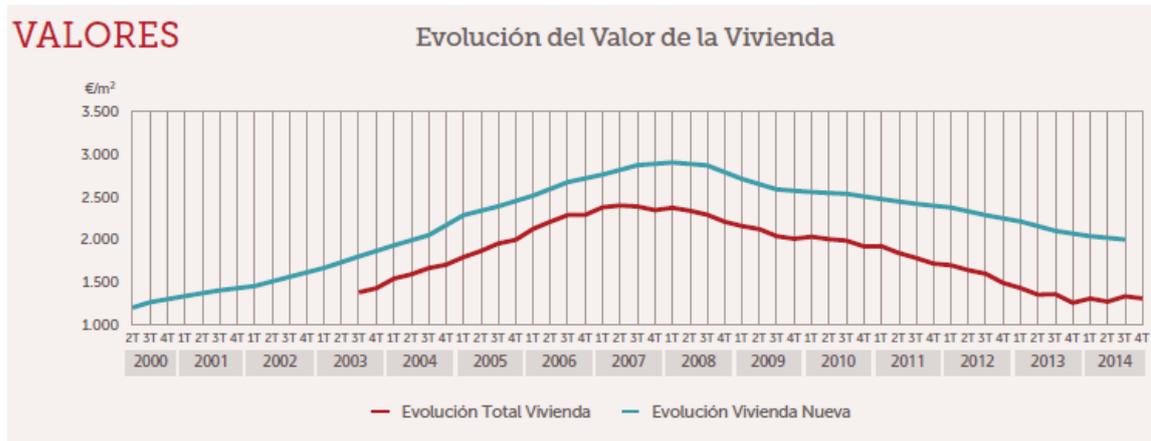
Esta burbuja, que acabó por pinchar a finales de 2007, hizo que los precios de la vivienda en España alcanzaran cotas nunca vistas hasta la fecha. Pero en este ámbito, no todas las zonas de España sufrieron con la misma intensidad los efectos de la crisis y del estallido de la burbuja. Mientras que las grandes ciudades presentaron en los barrios más céntricos una fuerte resistencia a la bajada de precios, los barrios más periféricos y las zonas alejadas de grandes núcleos de población (por ejemplo, las denominadas ciudades dormitorio) sufrieron severos ajustes en los precios por metro cuadrado. Del mismo modo, el mercado de segundas residencias, principalmente viviendas en las zonas de costa, vio desplomarse los precios durante más de siete años y todavía a día de hoy hay zonas donde los precios siguen bajando y no han tocado suelo todavía.

Si analizamos el caso de la costa cercana a Cartagena, observamos con datos actualizados que a fecha de hoy, mientras en zonas como Torrevieja el precio medio ya ha comenzado a subir, en las zonas aledañas al Mar Menor y en la propia ciudad de Cartagena los precios siguen experimentando un ajuste y todavía siguen bajando.



*1 Variación interanual de los precios de la vivienda en la zona costera de la provincia de Alicante y del Mar Menor; fuente: CartoDB y OpenStreetMap.*

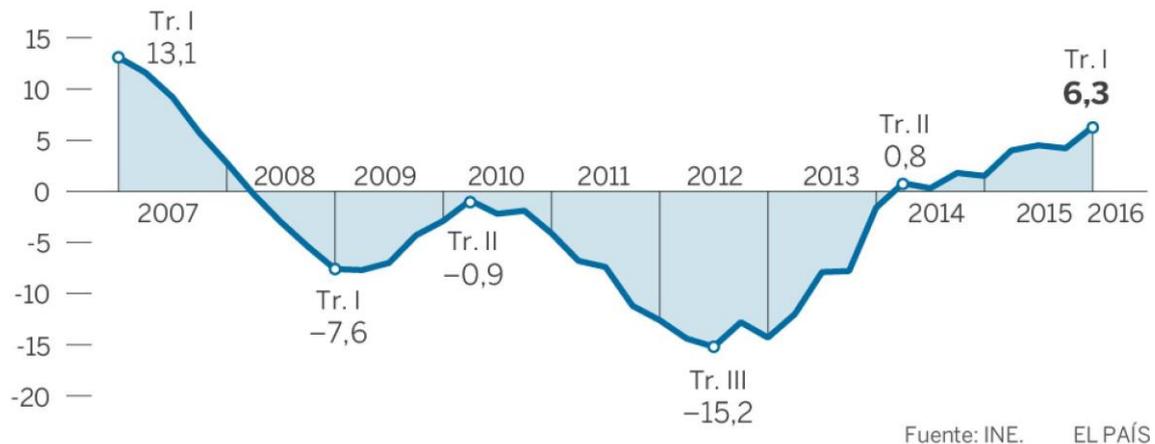
De este modo, considerando el año 2014 como punto de inflexión (en el año 2014 se vendieron un 22% más de viviendas que en 2013 según datos del Ministerio de Fomento), estamos viendo como diversos indicadores de actividad del sector se están reactivando, como la inversión extranjera en activos inmobiliarios (especialmente fuerte en grandes ciudades como Madrid) u otros indicadores como el consumo de cemento nos hablan de una cierta reactivación del sector de una manera lenta y moderada que ya está teniendo efecto sobre el precio medio de las viviendas.



2 Evolución del precio de la vivienda por metro cuadrado; fuente: web El Economista.

### PRECIO DE LA VIVIENDA

Variación interanual en %



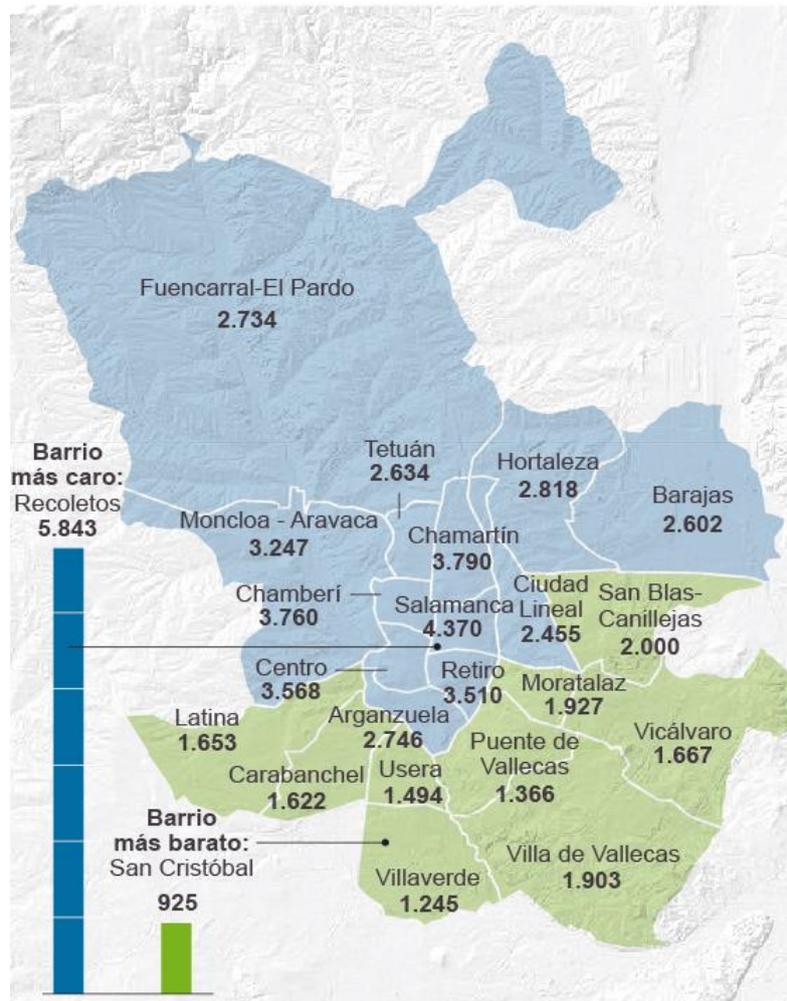
3 Variación interanual del precio de la vivienda; fuente: web EL PAIS con datos del INE.

## Evolución y análisis del precio de la vivienda en la ciudad de Madrid

El precio del metro cuadrado en Madrid capital (todo el municipio) se situó en el año 2015 en 2.663 €/m<sup>2</sup>, siendo la tercera capital de provincia más cara de toda España, sólo por detrás de San Sebastián y de Barcelona. Por distritos, Salamanca y Chamberí son los que presentan precios más elevados con 4.481 €/m<sup>2</sup> y 4.167 €/m<sup>2</sup> respectivamente. La burbuja inmobiliaria hizo que en la ciudad de Madrid, en el cuarto trimestre de 2008, el valor del precio medio por metro cuadrado en el distrito de Salamanca alcanzara su máximo histórico, ya que un piso en esta zona costaba de media 5.202 €/m<sup>2</sup>.

### PRECIO MEDIO DE LA VIVIENDA EN MADRID

■ Más de 2.000 euros / m<sup>2</sup>    ■ Menos de 2.000 euros / m<sup>2</sup>



Fuente: elaboración propia.

EL PAÍS

4 Precio medio de la vivienda en Madrid; fuente: web EL PAÍS.

El mapa del precio del metro cuadrado en el municipio de Madrid refleja de un modo claro las diferencias entre los distritos y barrios de la zona norte y los de la zona sur. Mientras que en la mitad norte los precios superan la barrera de los 2000 €/m<sup>2</sup> en la zona sur se quedan por debajo de esta barrera. La muestra estudiada en este trabajo cabe recordar que es la considerada como “almendra central” de Madrid, siendo los siete distritos que delimita el anillo de circunvalación de la M30, y que en el mapa se muestran como distritos de la zona en la que los precios medios superan los 2000 €/m<sup>2</sup>. Esta clara diferenciación norte-sur no ha hecho más que incrementarse tras la crisis y, en la actualidad, estas diferencias se están incrementando a medida que se implanta una leve mejoría en la economía del país. Así, desde el año 2014, año en el que se puede marcar el punto de inflexión en la tendencia de los precios de la vivienda, los distritos del municipio de Madrid con mayores incrementos en porcentaje de precio de mercado son los que se encuentran dentro del anillo de la M30 y que son el objeto de estudio en este trabajo: Centro, Salamanca, Retiro, Arganzuela, Chamartín, Chamberí y Tetuán.

Analizando los datos históricos durante los años de la crisis se observa cómo estos distritos son los que menos redujeron su precio durante los años 2007 a 2014 y cómo, además, fueron los primeros en comenzar a repuntar y a subir a partir del último trimestre de 2014. Este efecto es debido a múltiples factores, entre ellos que son los distritos con menos suelo disponible y donde menos promociones de vivienda nueva hay, de hecho esto es un fenómeno que se da en casi todas las grandes capitales y que se observa en España en grandes ciudades como Sevilla, Valencia o Barcelona, donde la vivienda en barrios céntricos sufrió los efectos de la crisis económica y el pinchazo de la burbuja inmobiliaria (bajada brusca de precios) de manera mucho más atenuada que otros barrios de la periferia.



*5 Variación anual por distritos en Madrid; fuente: web EL PAÍS.*

El cambio sustancial en la dinámica de caída de los precios en la ciudad de Madrid se apreció en el cambio de año de 2014 a 2015. Mientras en el año 2014 todos los distritos de Madrid capital cerraron con bajadas de precio interanuales, en el año 2015 ya hubo 12 distritos que presentaron subidas, todo ello en una ciudad con sus particularidades propias en cuanto a políticas de urbanismo y complejidad en cuanto a grandes operaciones que una gran capital presenta (operaciones Chamartín y Campamento, edificio España, etc.)

## Los ciclos del mercado de la vivienda



### 6 Reloj inmobiliario; fuente: La Caixa Research con datos del INE

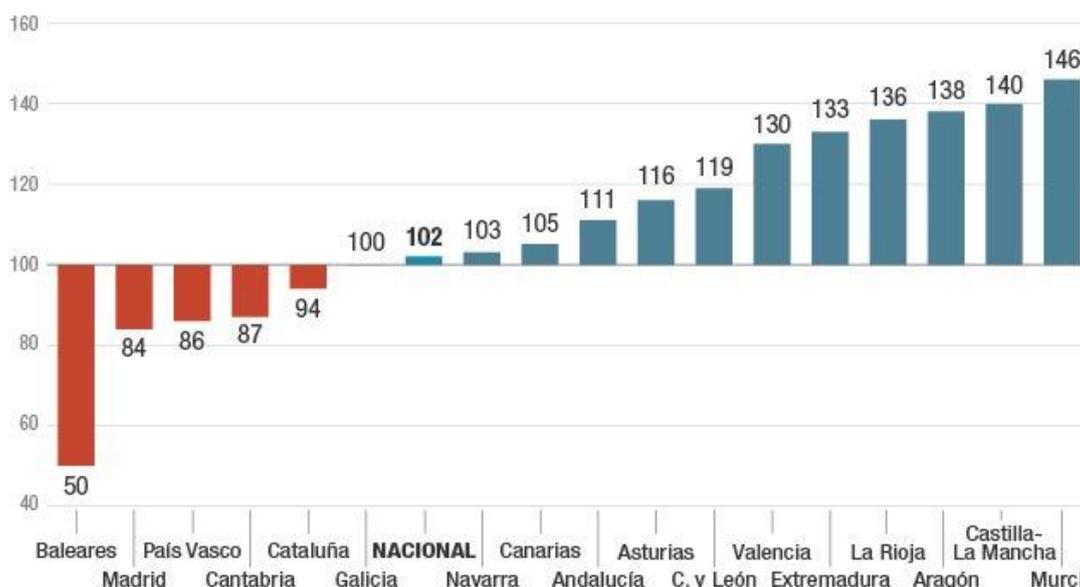
Al hablar del mercado inmobiliario es frecuente escuchar que se trata de un mercado *cíclico*, en el cual existen ciclos, al igual que en la economía, de expansión y crecimiento y otros de recesión o contracción del mercado. Del mismo modo, es común en los análisis del mercado inmobiliario la referencia a las *burbujas*, y ello es debido a que debido al peso que arrastra el sector de la construcción, cuando la economía crece es un sector que ostenta gran peso en el conjunto de la economía y que, del mismo modo, cuando este sector se ralentiza o decae, tiene grandes efectos de arrastre para la economía del país. En España, analizando los datos de la última década observamos como ese *reloj* del mercado inmobiliario llegó a sus cotas máximas de expansión y crecimiento en los años 2005 y 2006, en lo que se han venido a denominar los *años dorados* de la burbuja de la construcción. Después, en los años 2007 a 2009 comenzó un fuerte ajuste o desaceleración del sector, unido todo ello al comienzo de la gran crisis financiera global acaecida en todo el mundo y a la que España sumó los graves

efectos del estallido de la burbuja inmobiliaria. Se puede considerar que los efectos devastadores de la crisis y el ajuste del sector de la construcción duraron en España desde el año 2009 hasta el año 2013 o el 2014, años ambos en los que algunos indicadores del sector (precio del suelo, consumo de cemento, etc.) comenzaron a reflejar que la construcción había *tocado suelo* y el proceso de ajuste del sector había finalizado. En cualquier caso, es en el año 2015 cuando los expertos comienzan a hablar de algunos indicadores que marcan el nuevo comienzo de una etapa de expansión del sector, eso sí, muy lenta y moderada si se compara con otras etapas similares de décadas anteriores.

Por otro lado, no todas las ciudades y regiones experimentaron el estallido de la burbuja y sus consecuencias del mismo modo. Como hemos señalado anteriormente, el precio de la vivienda es algo muy dispar en las distintas regiones de España y la región de Madrid (y en concreto el municipio de Madrid) es una de las zonas más caras y de más difícil accesibilidad a la vivienda de todo el estado español (junto con País Vasco y Baleares). En el otro extremo, la Región de Murcia presenta la accesibilidad más fácil a la vivienda de todas las Comunidades Autónomas del país. En cualquier caso, indicadores como la compraventa de viviendas nos indican que el mercado inmobiliario se está estabilizando y comenzando a recuperarse, con la consiguiente subida de los precios, aunque no en todas las zonas de España por igual ni con la misma intensidad.

## Índice de accesibilidad para comprar una vivienda

Primer trimestre de 2016



El resultado con base 100 permite medir la distancia a la que se encuentran los salarios reales de los teóricos que son necesarios para comprar una vivienda por comunidades autónomas. Así, la media nacional da como resultado 102, es decir, que en términos globales los salarios reales que se cobran permiten endeudarse para adquirir casa. Pero persisten las diferencias por regiones. Si se superan los 100 puntos es más fácil y si no se llega, es más complicado.

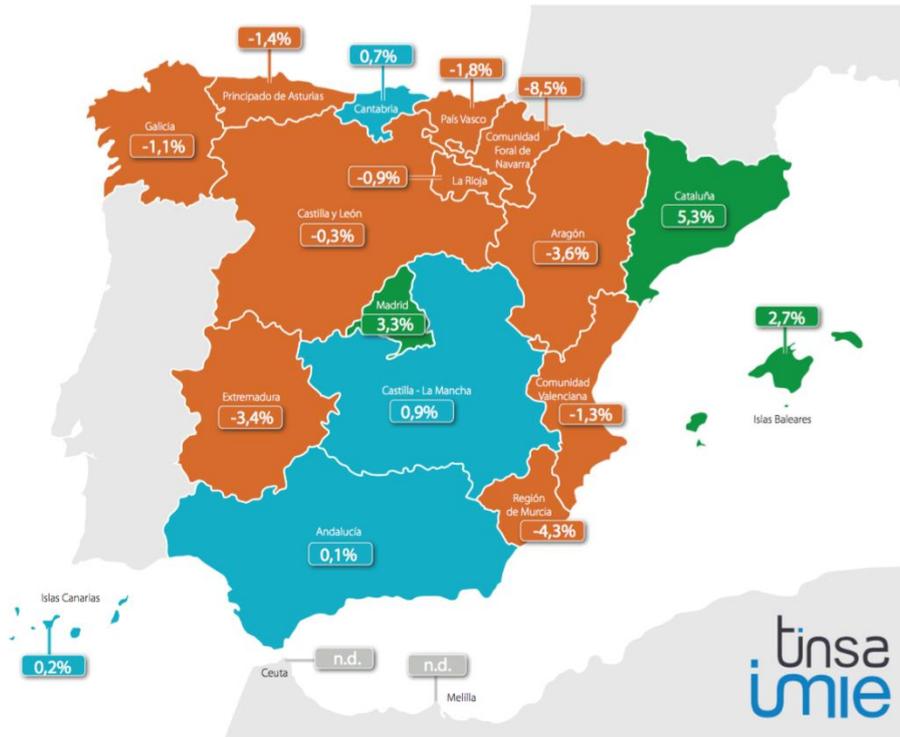
Fuente: Sociedad de Tasación

CARLOS CORTINAS / CINCO DÍAS

*7 Índice de accesibilidad para la compra de vivienda; fuente: Sociedad de Tasación.*

## MAPA RESUMEN DE LA VARIACIÓN INTERANUAL DEL PRECIO POR COMUNIDADES AUTÓNOMAS

Índice IMIE | Variación interanual



- Incremento superior a la media nacional
- Descenso superior a la media nacional
- Variación acorde a la media nacional (+/-1%)
- No disponible

### 8 Variación interanual del precio de la vivienda por Comunidades Autónomas; fuente: TINSA



### 9 Compraventa de viviendas 2010 - 2016; fuente: Instituto Nacional de Estadística

## El valor catastral y el precio de mercado

A la hora de valorar las viviendas o cualquier bien inmueble es frecuente manejar dos valores para cuantificarlo, el precio de mercado y el valor catastral. Mientras el valor catastral es un valor calculado con efectos administrativos y de control por parte de las administraciones competentes (generalmente los Ayuntamientos), el precio de mercado, al igual que ocurre con otros productos o servicios de muy distinta naturaleza, es el valor en el que según la teoría económica y en condiciones normales de mercado, la oferta iguala a la demanda y se produce la transacción entre comprador y vendedor.

En el presente trabajo se ha tratado la información correspondiente a los precios de mercado ya que la base de datos utilizada, extraída de un portal inmobiliario, aporta datos de precios de mercado de viviendas de la ciudad de Madrid. Por otro lado, y dado que es interesante el estudio y la relación entre ambos valores, hemos obtenido una base de datos de valores catastrales por barrio y hemos hecho un análisis de la relación de ambos a nivel de barrio. Con todo ello, las regresiones y el estudio en profundidad realizado en el trabajo se refiere siempre a los precios de mercado obtenidos de la citada web del portal *Idealista.com*.

El valor catastral es, por tanto, un valor que se fija objetivamente para cada inmueble y que resulta de la aplicación de unos criterios de valoración recogidos en la *ponencia de valores* de cada municipio. Es importante destacar que este valor no debiera superar el 50% del valor de mercado de la propiedad, y ello se puede comprobar en el estudio y el análisis realizado en el presente trabajo, así como en los mapas donde se refleja esta relación entre ambos valores. Así, el valor catastral de un bien inmueble se calcula en función de muchas variables, como su ubicación, su antigüedad, el coste de las obras realizadas en el inmueble, su estado de conservación, etc. Estos criterios de valoración son, según se recoge en la web oficial del catastro, los siguientes:

- La localización del inmueble y las circunstancias urbanísticas que afecten al suelo (valor de mercado del suelo).
- El coste de ejecución material de las construcciones (valor de mercado de la construcción), los beneficios de la contrata, honorarios profesionales y tributos que gravan la construcción, el uso, la calidad y la antigüedad edificatoria, así como el carácter histórico-artístico u otras condiciones de las edificaciones.
- Los gastos de producción y beneficios de la actividad empresarial de promoción.

El valor catastral de los inmuebles no podrá superar el valor de mercado y es un valor utilizado como base para el cálculo de impuestos afectos a la propiedad. Del mismo modo, los valores catastrales se actualizan anualmente mediante la aplicación de coeficientes que se aprueban en las Leyes de Presupuestos Generales del Estado (fuente: <http://www.catastro.meh.es>).

A la hora del cálculo de impuestos hay que tener en cuenta sobre qué valor se calculan, ya que algunos de ellos se calculan en base al valor catastral de la vivienda, como es el caso del IBI (Impuesto local de bienes inmuebles) y otros, en cambio, se calculan en base al precio de venta o de mercado, como es el caso del ITP (Impuesto de transmisiones patrimoniales). Éste último,

que suele rondar el 7% del precio de la venta, en determinadas situaciones (como ha venido ocurriendo en tiempos de crisis donde muchas viviendas se venden por el importe de hipoteca que queda pendiente de satisfacer), puede llevar a malentendidos y desacuerdos con Hacienda debido a que las comunidades autónomas tienen unas tablas de precios mínimos sobre las que calculan el *ITP* más bajo que debe pagar el comprador del inmueble. Esta práctica trata de evitar el fraude que supone escriturar un inmueble por un valor inferior al real, y se ha dado con frecuencia en los casos de venta de viviendas por muy bajo valor (los *chollos*). Por todo ello, es conveniente tener conocimiento de las distintas maneras de valorizar un bien inmueble y conocer cuál de ellas (valor catastral o precio de mercado) se utiliza como base a la hora del cálculo de los distintos tributos afectos a la compraventa y posesión de bienes inmuebles.

Es importante también señalar que, a diferencia del precio de mercado (que varía y fluctúa con las condiciones reales del mercado inmobiliario en épocas de crecimiento y recesión), el valor catastral no cambia con las condiciones del mercado. Esto es así por su propia naturaleza, ya que el valor catastral se calcula en función de numerosas variables pero no está sujeto a las condiciones económicas coyunturales.

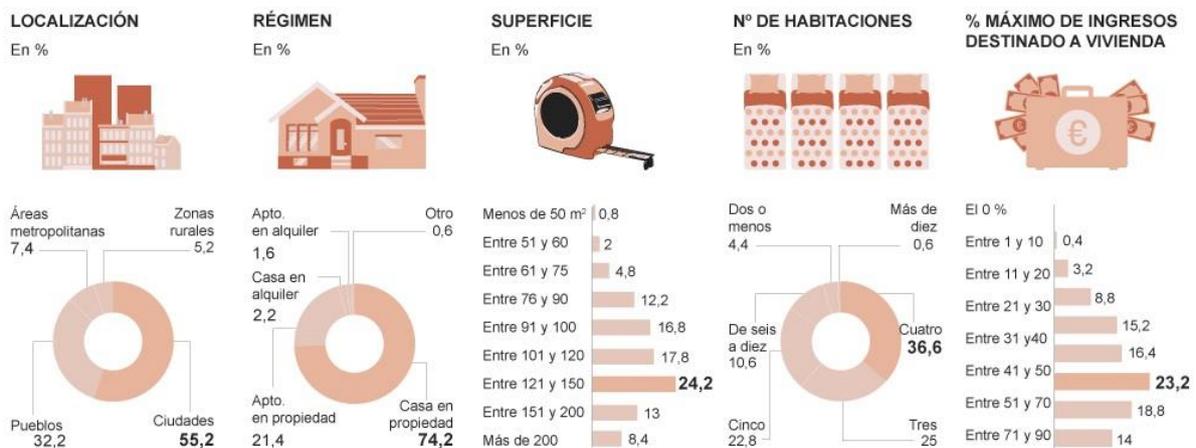
## Factores determinantes del precio de la vivienda

El precio de la vivienda se encuentra afecto a factores internos y externos. Los factores internos (*indoor*) son aquéllos relativos a las características propias de la vivienda, tales como su tamaño, su orientación, el número de habitaciones o de baños que disponga, el estado de conservación de la construcción (si es nueva, de segunda mano, rehabilitada...) y otros tantos que describan cualidades propias del inmueble. Estos factores internos no comprenden todos aquéllos que tienen en cuenta la localización espacial de la vivienda.

En consecuencia, los llamados factores externos (*outdoor*) son los que describen a la vivienda en función de su localización espacial. Entre estos factores externos podemos destacar los relativos a la ubicación física de la propia vivienda, tales como la cercanía o la lejanía a determinadas zonas o puntos de interés como puedan ser, por ejemplo, estaciones de metro, centros comerciales, parkings públicos, colegios, universidades, museos, hospitales, zonas verdes o de recreo, etc.

En cuanto a los factores *indoor* de una vivienda, las características principales analizadas por un potencial comprador de vivienda son la localización del inmueble (zona urbana, rural, casco urbano, extrarradio...), la superficie del inmueble y el número de estancias (baños y habitaciones principalmente).

## Así es la vivienda más deseada



10 Características de la vivienda deseada por los españoles; fuente: [elmundo.es](http://elmundo.es)

## JUSTIFICACIÓN DEL TRABAJO

### Objetivos y propósito de la investigación

Este trabajo fin de máster (en adelante, TFM) pretende realizar una investigación acerca de la formación del precio de la vivienda en la ciudad de Madrid desde una perspectiva espacial. Los precios de la vivienda tienden a estudiarse en numerosas investigaciones desde un punto de vista temporal, analizando los ciclos en los que los precios suben y en los que bajan o se mantienen. En cambio, son menos los trabajos que analizan el fenómeno desde el enfoque de la econometría espacial, teniendo en cuenta variables y factores como los que en este trabajo se analizan.

Los pilares en los que se basa el presente trabajo de investigación son varios. Por un lado se ha obtenido una base de datos real y actualizada con 4.100 referencias de viviendas en venta en la zona analizada de la ciudad de Madrid. En dicha base de datos se cuenta con información acerca de las características propias de dichas viviendas, como su tamaño, número de habitaciones, etc. Con ello tenemos información suficiente para realizar una aproximación a la formación del precio de la vivienda desde un modelo de regresión hedónica, que explica la importancia de cada factor o variable en la formación de dicho precio. Por otro, aprovechando que la base de datos obtenida nos ofrece información acerca de la georreferenciación de cada vivienda (su ubicación mediante la latitud y la longitud), se optó por hacer un estudio espacial, tanto a nivel de modelos de regresión espacial como mediante la utilización de sistemas de información geográfica (SIG) con los que dar soporte a la investigación.

Para el estudio espacial se ha estudiado el factor de lejanía/cercanía a diversos elementos de la ciudad para analizar la influencia sobre la variable precio, a saber, cercanía a zonas verdes, colegios, hospitales, etc. Mediante una exhaustiva búsqueda de elementos de la ciudad de Madrid georreferenciados y el cálculo de distancias mínimas a todas las observaciones de la base de datos se ha analizado mediante modelos de regresión la importancia de cada factor en la variable dependiente, que es siempre el precio de la vivienda en cuestión. Mediante el uso de SIG, se pretende dar al trabajo una orientación que, además de la pura estadística utilizada en los análisis de regresión, ayude a comprender de un modo gráfico la estructura espacial de una variable en una determinada zona urbana como la analizada.

Por último, este trabajo pretende ahondar en el estudio de las variables con efectos espaciales en el estudio econométrico, ya que cada vez más se incorporan estos factores al estudio y al análisis socioeconómico, pues ayudan a crear modelos que permiten comprender de una manera adecuada los entornos y fenómenos analizados.

## ESTADO DE LA CUESTIÓN

Las investigaciones realizadas en el campo de la econometría espacial donde se utilizan modelos de regresión que toman en cuenta factores espaciales comienzan a generalizarse en la década de los 90, principalmente desde los estudios de **Luc Anselin** y otros investigadores que, bajo su influencia, comenzaron a tener en cuenta los factores espaciales en sus estudios econométricos en Estados Unidos. Anselin es actualmente el director de la *School of Geographical Sciences and Urban Planning* de Chicago y el creador del software **GeoDA**, por tanto se le considera pionero en los estudios de econometría espacial y en los estudios geoestadísticos. Además de Anselin, los otros dos investigadores destacados a nivel mundial en análisis espacial aplicado a las ciencias sociales son **Stewart Fotheringham** (director del *Centre for GeoInformatics de la Universidad de St. Andrews*), y **Michael Batty** (director del *Centre for Advanced Spatial Analysis de la Universidad de University College*).

A lo largo de los 90 y en la década 2000- 2010 se continuó con la senda de los trabajos científicos que aportaban nuevas visiones acerca de los modelos de regresión que tienen en cuenta estos factores (GWR) y los estudios de diversas materias bajo el paraguas de la econometría espacial (estudios sobre salud y enfermedades, medioambiente o precios de la vivienda entre otros). Además, con los métodos de precios hedónicos y mediante regresiones hedónicas se han analizado en profundidad campos como el que nos ocupa en este trabajo, el de la formación de los precios de la vivienda.

Hay que destacar que, con frecuencia, en la literatura internacional se habla de análisis espacial o de las técnicas de la econometría espacial utilizando el término *espacial* como abreviación de *espaciotemporal* (Anselin, 1995; Fotheringham et al., 1998; Granis, 2002). Los campos del conocimiento donde tiene aplicación el análisis y las técnicas de econometría espacial son muy diversos, tanto en las ciencias de la materia (física, química), como en las ciencias de la vida (salud, biología, ecología) y en las ciencias sociales y las humanidades (desarrollo urbano, economía, sociología, historia, etc).

Es, sin embargo, desde el año 2000 y hasta hoy, y con el auge y el perfeccionamiento de los SIG, cuando se han publicado trabajos apoyados en los potentes softwares SIG y de cálculo estadístico que aún a día de hoy continúan demostrando la importancia de incluir factores espaciales en el estudio de numerosos fenómenos socioeconómicos. Centrándonos en el ámbito del estudio del precio de la vivienda, algunos de los estudios y artículos estudiados para la elaboración de este TFM, ordenados cronológicamente, serían los siguientes:

En el año 2000, los trabajos de investigación de Tyrväinen y Mettinen muestran que en Salo (Finlandia) se paga un 4,8 % más por una vivienda que tenga vistas a un bosque. Asimismo, la proximidad a un parque forestal también afecta al precio de la vivienda, cada kilómetro adicional que esté más alejada reduce el precio de ésta en un 5,7 %. Estos estudios están claramente en la línea de lo realizado en este TFM, con enfoque específico en los parques y en las zonas verdes en general. En la línea de los estudios de impacto medioambiental, también en el año 2000 Wilhelmsson analiza el impacto que tiene el ruido procedente del tráfico rodado sobre el valor de viviendas unifamiliares en un barrio de Estocolmo (Suecia). En particular, los resultados obtenidos demuestran que, en promedio, por

cada decibelio adicional el precio de la vivienda se reduce un 0,6%, mientras que una casa situada en un lugar ruidoso vale, en promedio, un 30% menos que otra situada en un lugar tranquilo. Este tipo de estudios refleja la tremenda complejidad del análisis del precio de la vivienda desde la perspectiva hedónica, ya que son muy numerosos y casi infinitos todos los aspectos y variables explicativas que un potencial comprador puede llegar a considerar a la hora de comprar el bien que por la mayoría de las personas es considerado de mayor importancia a lo largo de su vida. No solo el tamaño o el número de habitaciones o baños, sino factores espaciales como cercanía a determinados elementos o zonas con mayor o menor ruido, mejores o peores vistas, o incluso mejor o peor clima dentro de una misma área urbana pueden determinar la variable del precio de una vivienda.

Coetáneos a estos estudios son los trabajos de Bond y Seiler en el año 2002, que examinan el efecto que las vistas del lago Erie en Cleveland (EE.UU) tienen sobre el valor de una casa. En sus trabajos señalan que una de las situaciones más agradables en esta vida es la sensación de placer que proporciona la vista de una gran extensión de agua. En su estudio, muestran que una casa con vistas al lago cuesta 256.544 dólares norteamericanos (un 89,9%) más que una vivienda media que careciera de esta característica. Las dificultades de estos trabajos radican en la precisión a la hora de calcular lo agradable o desagradable que resulta la visión de un lago para las personas, ya que es algo que de un territorio a otro puede variar sensiblemente y hace que estos estudios tengan una validez más local que universal.

El tema del impacto visual, en este caso en su variante negativa también se estudia en el año 2003 por el investigador Des Rosiers, que analiza el impacto que los tendidos eléctricos de alta tensión tienen sobre el precio de las viviendas en Brossard, cerca de Montreal (Canadá). En sus estudios concluye que el impacto visual que tiene la vista directa de una torre de alta tensión ejerce un impacto negativo sobre el valor de la vivienda con una reducción del valor de la misma de aproximadamente un 10%.

En España, en el período que va desde el año 2000 hasta la actualidad se han desarrollado varias tesis doctorales sobre aspectos relacionados con el análisis del precio de la vivienda basado en factores espaciales, algunos de ellos más teóricos mediante análisis estadísticos y regresiones y otros, sobretodo últimamente, que utilizan los SIG para apoyar sus resultados y conclusiones mediante el análisis geográfico y visual de los datos.

Las tesis doctorales de **Mónica Navarrete Álvarez** de 2011 titulada *Modelos geoestadísticos del precio de la vivienda: aproximación al conocimiento intraurbano de la ciudad de Madrid* es un estudio en profundidad de la estructura espacial de la ciudad de Madrid y cuyo objetivo es *la aproximación al conocimiento del fenómeno urbano en las distintas escalas del territorio*, algo para lo que se apoya en la econometría espacial y en el estudio de los factores espaciales que afectan al precio de la vivienda, lo cual trata desde el punto de vista de los modelos geoestadísticos.

Por otro lado, la directora de la tesis de Mónica Navarrete, la profesora de la Universidad Autónoma de Madrid (UAM) **Coro Chasco Yrigoyen**, en su tesis doctoral de 2003 titulada *Econometría espacial aplicada a la predicción-extrapolación de datos microterritoriales* analiza datos microterritoriales con modelos basados en los postulados de la econometría espacial y en base al análisis exploratorio de datos espaciales (AEDE). Los apuntes de las asignaturas que

imparte en sus clases de la Universidad Autónoma de Madrid sobre el AEDE han sido utilizados en el estudio y análisis de los datos y resultados de este TFM.

Los investigadores de la Universidad de Castilla-La Mancha **José María Montero, Beatriz Larraz, Gema Fernández-Avilés y Matías Gámez** han venido trabajando en los últimos años en los campos de la geoestadística y la econometría espacial aplicada al sector de la vivienda con diversos trabajos que exponen diversas metodologías en cuanto a las regresiones espaciales y a la utilización de Sistemas de Información Geográfica (SIG). En sus trabajos *“Estimación espacial del precio de la vivienda mediante métodos de Krigeado”*, *“Una aplicación de la estadística espacial al comportamiento de la vivienda de alquiler en España”* y *“Aportaciones de los métodos geo-cuantitativos a la estimación de precios de vivienda”* exponen los resultados que confirman la importancia de los factores y de las variables explicativas con efectos espaciales, así como la confirmación de que la existencia de autocorrelaciones espaciales en la formación de los precios de la vivienda es clara y evidente. Algunas de las conclusiones de estas investigaciones recientes son que *“los Sistemas de Información Geográfica (SIG) nos proporciona el soporte ideal para realizar estudios de estadística espacial, es decir, para explicar fenómenos que muestran un patrón de comportamiento dependiente del espacial”*, que *“no se puede obviar el efecto del espacio en los estudios de ámbito regional, como es el caso de la formación de los precios de la vivienda”* y, por último, que *“parece evidente la importancia del conocimiento de los precios de vivienda en un contexto como el actual en el que el nivel de actividad de los agentes económicos participantes en este mercado es muy intenso. Por consiguiente, en una tesitura como la actual, de poco valen los precios medios a escala provincial o de Comunidad Autónoma publicados trimestralmente por el Ministerio de Fomento o suministrados con cierta regularidad por determinadas entidades tasadoras. A todos los efectos, sería más provechoso para los agentes económicos participantes en este mercado disponer de estimaciones en cualquier punto del área objeto de estudio. Ello permitiría a un futuro comprador o vendedor hacerse una idea más real del precio que tendría que pagar (recibir) por un inmueble de determinadas características ubicado en un determinado punto de una zona geográfica”*.

Esta última conclusión redundante en la idea de que los datos sobre precios de viviendas agregados a nivel de Comunidad Autónoma o de municipio no son del todo útiles para el potencial comprador, y por ello se hace preciso y necesario el estudio de modelos que sepan predecir de forma exacta y con precisión la cuantificación del precio de un inmueble atendiendo a sus características y a los factores espaciales de su entorno.

También son numerosos los artículos de investigadores chinos en los que, haciendo uso de los SIG y aplicando regresiones espaciales (GWR), se analizan los efectos espaciales que sobre el precio de la vivienda tienen diferentes elementos en distintas ciudades, como es el caso de la ciudad de Shenzhen en el artículo *“Geographically weighted regression model (GWR) based spatial analysis of house price in Shenzhen”* y en cuyo abstract, en inglés, leemos un resumen de lo que podría ser este TFM aplicado a una ciudad de China: *Through applying spatial statistical analysis, **Geographical Weighted Regression (GWR) model and GIS technology**, this study aims at **finding the relationship between the effects of various factors and spatial distribution of residential house price**. The traditional regression models are reviewed firstly, the model without the consideration of spatial characteristics cannot reach very nice precision*

*to simulate the spatial distribution of the house price. In this study, the spatial statistical model, coupled with GIS as well as GWR model, is developed.*

Otros ejemplos de investigaciones chinas acerca del precio de la vivienda son las publicadas en los artículos *Influence of Spatial Features on Land and Housing Prices* de Gao Xiaolu, *Spatial Pattern and Structural Determinants of Shanghai's Housing Price: a GWR-based approach* de Qingyuan Tang y *Spatial econometric analysis of housing price of chinese provinces* de Hu Zhaohui. En todos ellos se hace un estudio de la influencia de factores espaciales y su influencia en los precios de la vivienda en diferentes regiones y ciudades chinas con modelos econométricos basados en regresiones GWR (espaciales).

En España, más recientemente, otro de los ejemplos de aplicación de la econometría espacial aplicado además al análisis de la ciudad de Madrid es el artículo *"The application of spatial analysis to cadastral zoning of urban areas: an example in the city of Madrid"* de **María del Carmen Morillo, Francisco García Cepeda y Sandra Martínez-Cuevas**. Este trabajo, del año 2016 y recién publicado, sostiene en una de sus conclusiones que *"one of the most important conclusions is the validity of the use of these interpolation techniques along with GIS techniques to evaluate spatial correlations as Tobler's Law [Tobler, 1970]"*. Dicho de otro modo, las últimas investigaciones se centran en exponer que, mediante el uso de los SIG, se corrobora la existencia de autocorrelación espacial en el fenómeno del precio de la vivienda y, por ende, se confirma lo expuesto en la Ley de Tobler, ya citada en este TFM y que afirma que *"todas las cosas están relacionadas con todo lo demás, pero las cosas cercanas entre sí lo están más que las distantes"*.

El último de los trabajos analizados, por ser el más reciente, es el trabajo de los ya mencionados investigadores del Departamento de Economía y Empresa de la Facultad de Ciencias Jurídicas y Sociales de la Universidad de Castilla-La Mancha en Toledo José María Montero Lorenzo y Gema Fernández Avilés titulado *"La importancia de los efectos espaciales en la predicción del precio de la vivienda. Una aplicación geoestadística en España"* en el que hacen hincapié en la importancia de incluir los factores espaciales en el estudio econométrico del precio de la vivienda y demuestran la influencia de estos factores en la formación de la variable dependiente del precio, muy en la línea de lo trabajado en este TFM y con similares conclusiones. Las líneas de los últimos trabajos de investigación van encaminadas a confirmar la existencia de autocorrelación espacial en la formación de los precios de la vivienda y a confirmar la necesidad de tener en cuenta los factores espaciales a la hora de elaborar modelos econométricos para predecir o calcular precios de viviendas.

Por último, no sólo en los artículos científicos y académicos se exponen los trabajos de investigación realizados en este campo. Es frecuente ver en prensa, siempre acompañados de mapas temáticos y gráficos explicativos, noticias como la adjunta, elaborada por una consultora especializada en estudios urbanos, donde se analiza la influencia de la cercanía a colegios de la ciudad de Madrid sobre el precio de las viviendas. Estudios de este tipo, cuando se centran exclusivamente en el análisis de una de las variables espaciales (en este caso los colegios), no son demasiado concluyentes (más allá de la búsqueda del titular fácil y que busca la llamada de atención del lector), ya que, como opinan los propios expertos del sector inmobiliario, no se puede reducir el estudio del precio de la vivienda a la influencia de un único

factor debido a la complejidad del mismo, tratándose de uno de los fenómenos socioeconómicos más interesantes y más estudiados en los últimos tiempos debido a fenómenos como el estallido de la burbuja inmobiliaria en España. En el siguiente extracto de la noticia se hace alusión a las técnicas de *big data* (tratamiento de ingentes cantidades de volúmenes de datos) y a la imposibilidad de asociar el precio de la vivienda directamente con un determinado factor únicamente: *¿Influyen estos factores en el precio de las casas? Pese a que todos estos expertos en el mercado de la vivienda en Madrid reconocen la importancia de la oferta educativa y su estrecha relación con el mercado residencial, ninguno cree que se pueda cuantificar en qué medida la proximidad de un colegio sea determinante a la hora de fijar el precio de las viviendas. "No sería correcto generalizar", manifiesta Gil. "Es una variable difícil de cuantificar, salvo en nuevos barrios en los que el efecto llamada sería inmediato", explica López Van-Dam. Y no les falta razón, puesto que, pese a ser muy importante para las familias, el colegio no deja de ser uno más de los factores de decisión a la hora de escoger un lugar de residencia para el que también pesan mucho otros servicios, como el transporte público. Por lo tanto, la relación directa con el mercado de la vivienda es muy difícil de medir. Pero no imposible. Actualmente, y gracias al empleo de las técnicas de big data, puede saberse en qué medida afecta la oferta escolar de una zona al precio de las viviendas cercanas. Para ello, este suplemento ha contado una vez más con la colaboración de Urban Data Analytics (uDA), firma especializada en el análisis de datos relacionados con el mercado inmobiliario.*

SU VIVIENDA • Análisis inmobiliario de la ciudad de Madrid con tecnología 'big data'

## Las casas a menos de 500 metros de un colegio son un 6% más caras



GRÁFICO INTERACTIVO: uDA (urban Data Analytics)

- Las viviendas situadas cerca de centros concertados valen un 9% más que las que están próximas a los públicos. Junto a los privados suben un 28%
- Así afecta el Metro al precio de la vivienda en Madrid

**11 Noticia del 11 de Septiembre de 2016 sobre la influencia sobre el precio de la vivienda de la cercanía a colegios en la ciudad de Madrid; fuente: [elmundo.es](http://elmundo.es)**

## METODOLOGÍA

### La econometría espacial

Para explicar la metodología utilizada en la elaboración del presente TFM es preciso explicar en qué consiste la disciplina de la econometría espacial.

Es incuestionable la importancia del espacio dentro del estudio estadístico de variables de carácter socio económico, y por ello, especialmente en los últimos años, está tomando gran relevancia en los estudios de econometría la econometría espacial, que no es otra cosa que la parte de la econometría que se dedica al estudio de los fenómenos económicos espaciales.

Aunque todas las actividades económicas se desarrollan en un espacio determinado, por fenómenos económicos espaciales se entienden aquellos en los que la variable espacio juega un rol tan importante que su exclusión podría dar lugar a modelos econométricos con severos errores de especificación.

La importancia del espacio, se manifiesta a menudo en la autocorrelación espacial, y según *Paelink y Klaassen (1979)* existen cinco principios fundamentales que rigen al análisis espacial:

- Interdependencia: la dependencia recíproca entre las diferentes unidades de análisis deben ser debidamente incorporadas.
- Asimetría: refleja la idea de concentración y desconcentración gradual en diferentes áreas.
- Alotopía: la causa de un fenómeno espacial debe buscarse en otro lugar.
- No linealidad.
- Topología: Inclusión de variables de distancias entre dos localizaciones, coordenadas, densidades y otras.

Muchos economistas se han resistido a la econometría espacial por considerar que las variables espaciales se adicionan al análisis sólo por mostrar significatividad estadística sin ninguna justificación teórica, pero los estudios espaciales dentro de la corriente principal de la economía han ido incorporando variables de tipo espacial y cada vez más se consideran las variables espaciales en los análisis estadísticos y en los modelos econométricos. De hecho, existe una gran cantidad de ramas de la economía que han incorporado a sus análisis variables de carácter espacial. Esto es debido también, en parte, al desarrollo de potentes herramientas informáticas y de softwares que facilitan el estudio de variables espaciales que de otro modo resultarían más difíciles de analizar. Algunas de estas ramas pueden ser la economía urbana o la economía regional (con ejemplos como el tratado en este TFM, precios de vivienda, gasto público de administraciones, economía de la salud en determinadas regiones, incidencia de enfermedades por zonas, etc.). En este sentido, uno de los aportes teóricos más significativos

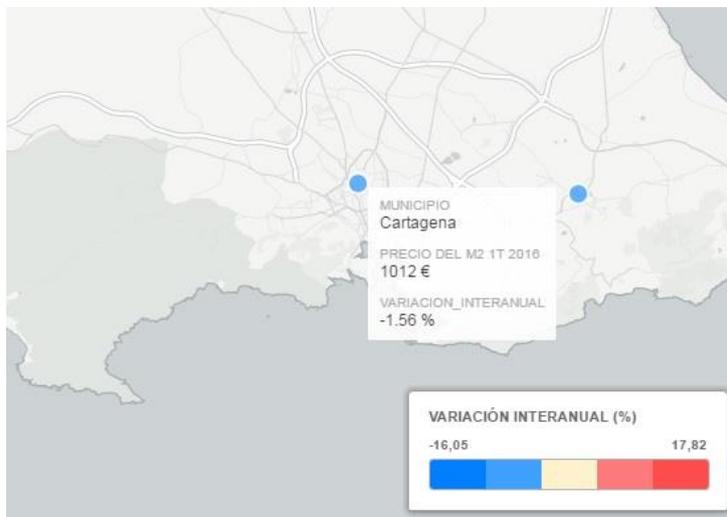
provino de la Teoría del Desarrollo Económico y de la denominada *Teoría de la Organización Espacial* (Posada, 1978).

Habitualmente, en el estudio econométrico de fenómenos económico-espaciales suelen surgir dos problemas, que son inherentes a la propia naturaleza de la estructura de datos a analizar, a saber, la heterogeneidad espacial y la autocorrelación espacial (también llamada dependencia espacial por algunos autores). Para el análisis de la autocorrelación espacial, en este trabajo se han utilizado contrastes como el índice de Moran o los mapas de significancia LISA para la correlación local.



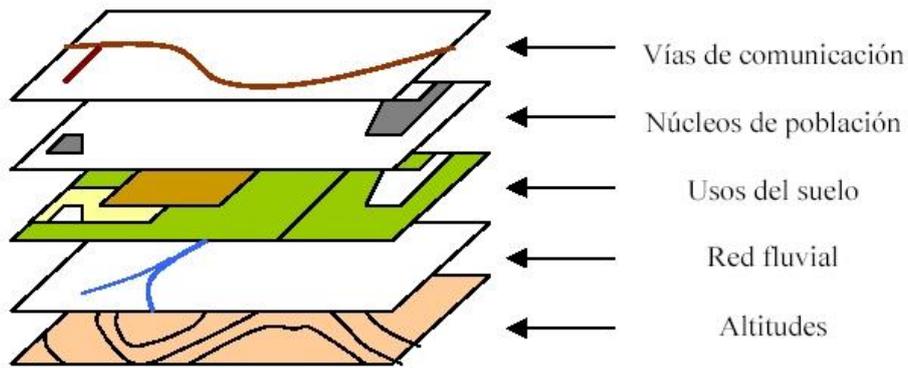
Fuente: Tinsa

**12 Variación interanual del precio de la vivienda en la costa española; fuente: OpenStreetMap y CartoDB con datos de TINSA.**



**13 Variación interanual del precio de la vivienda en Cartagena; fuente: OpenStreetMap y CartoDB con datos de TINSA.**





*15 Un SIG puede mostrar la información en capas temáticas para realizar análisis multicriterio complejos.*

La georreferenciación queda definida por una función matemática del tipo:

$$X= f(x, y)$$

$$Y= f(x, y)$$

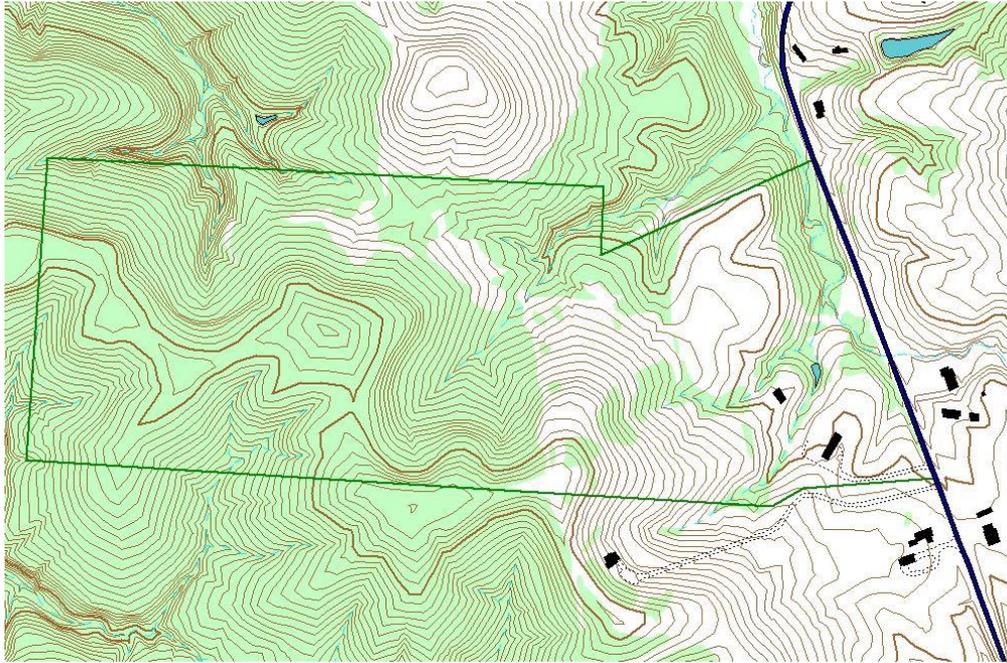
Donde la posición de una entidad geográfica en el sistema de coordenadas destino (X, Y) es función de las coordenadas (x, y) que tiene ese elemento en el sistema origen.

Para poder realizar una georreferenciación es necesario identificar sin lugar a equivocación puntos homólogos en los sistemas de coordenadas de origen y destino, lo que permite calcular los parámetros de la transformación. Algunos de los factores que afectan a la calidad de la rectificación son el número de puntos homólogos identificados y la distribución de estos puntos en la superficie del mapa.

La georreferenciación es, por tanto, el uso de coordenadas de mapa para asignar una ubicación espacial a entidades cartográficas.

El método más utilizado para describir la posición de una ubicación geográfica en la superficie de la Tierra consiste en utilizar mediciones esféricas de latitud y longitud. Estas son mediciones de los ángulos (en grados) desde el centro de la Tierra hasta un punto en su superficie. Este tipo de sistema de referencia de coordenadas generalmente se denomina sistema de coordenadas geográficas. Así, en los mapas generados se superponen capas con los distintos elementos georreferenciados para de este modo analizar relaciones espaciales que puedan existir entre ellos.

Para georreferenciar elementos, como en el caso de este trabajo donde se ha llevado a cabo la georreferenciación de las viviendas estudiadas y los diversos elementos analizados (parques, colegios, hospitales...) hacemos uso de los SIG, que no dejan de ser más que software con herramientas de georreferenciación.



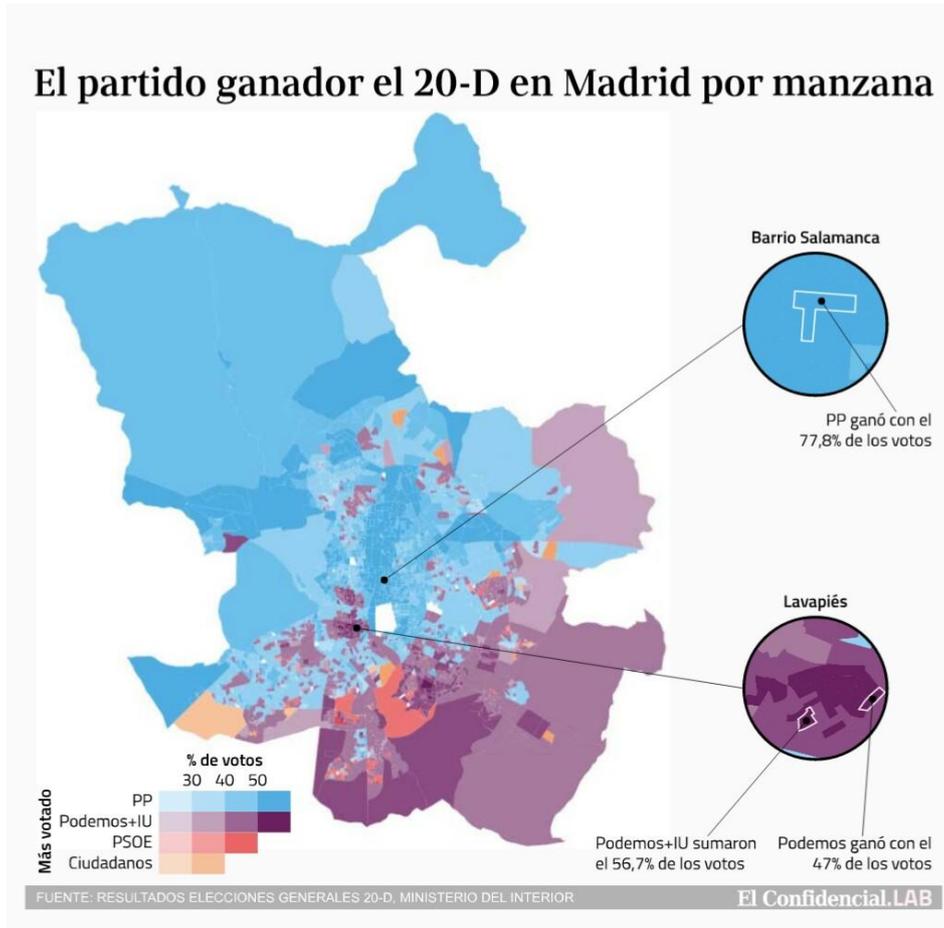
*16 Un ejemplo de uso de la superposición de capas en una aplicación SIG*

En el ámbito de las herramientas SIG existen diferentes aplicaciones, tanto comerciales como libres, **GeoDa**, **ArcGis** (el utilizado en este TFM) o **SaTScan** son solo algunos de ellos. En la última década el propio concepto de georreferenciación y el uso de los SIG en diferentes ámbitos se ha difundido más allá del ámbito científico-tecnológico gracias a herramientas como **Google Earth**, cuya facilidad de uso ha extendido y democratizado el tratamiento de información de mapa fuera del ámbito técnico existente hasta ahora, haciéndolo accesible al público no científico.

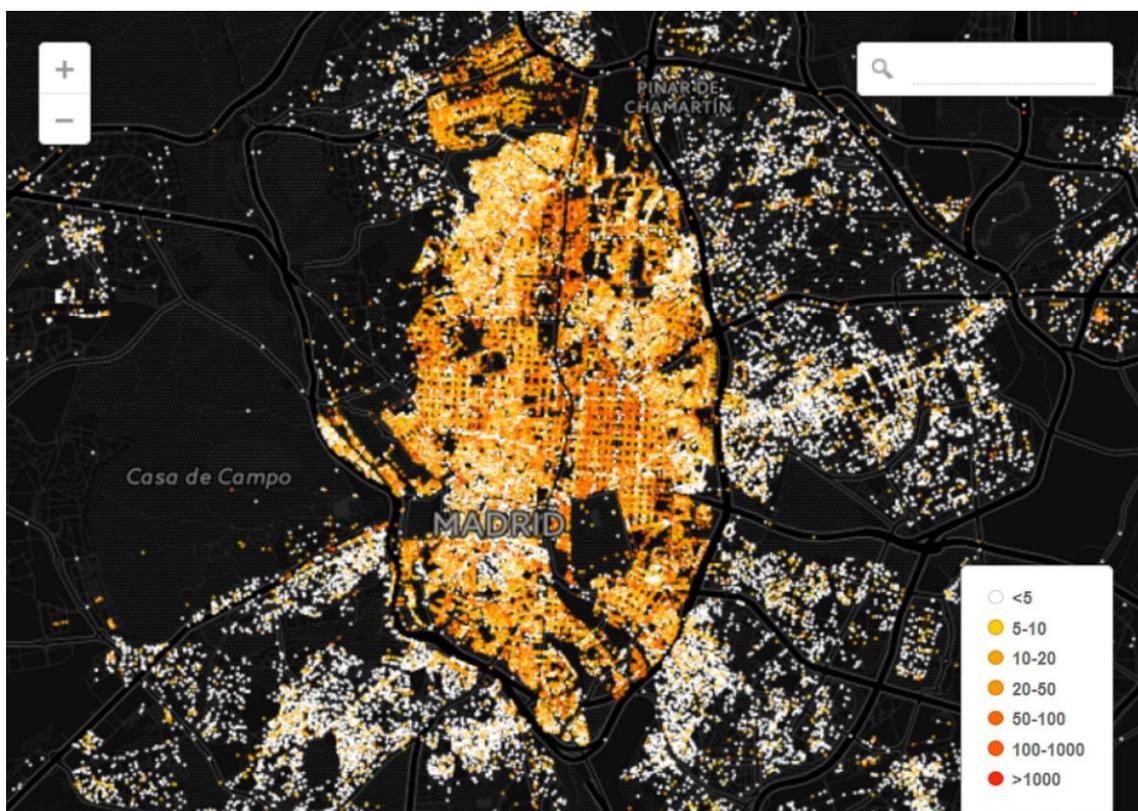
El software utilizado en esta investigación es **ArcGIS** (en su versión online) de **ESRI**, una aplicación que, según la definición que aparece en su web comercial, *conecta mapas, aplicaciones, datos y personas para que puedas tomar decisiones más inteligentes y rápidas, proporcionando a todos los miembros de una organización la capacidad de crear, analizar, almacenar y compartir mapas desde cualquier dispositivo, en cualquier lugar y momento.*

La plataforma **ArcGIS** pone a disposición de los usuarios contenidos entre los que se incluyen mapas base o imágenes que se adaptan a las necesidades de cada proyecto u organización. Por ejemplo, en esta investigación se han utilizado mapas base y capas de la ciudad de Madrid, así como elementos georreferenciados disponibles de manera abierta y gratuita (creados por otros usuarios) como la red de líneas y estaciones de metro. En los mapas creados en esta aplicación hemos georreferenciado todos los elementos a analizar en el trabajo para visualizar los resultados que se han evaluado mediante modelos de regresión y así poder constatar de manera visual la estructura espacial de la vivienda en la zona analizada, además de aprovechar las funciones de análisis de datos que el programa ofrece, y que se pueden ver y analizar en los mapas que aparecen en la sección de *Resultados* del presente estudio.

## El partido ganador el 20-D en Madrid por manzana



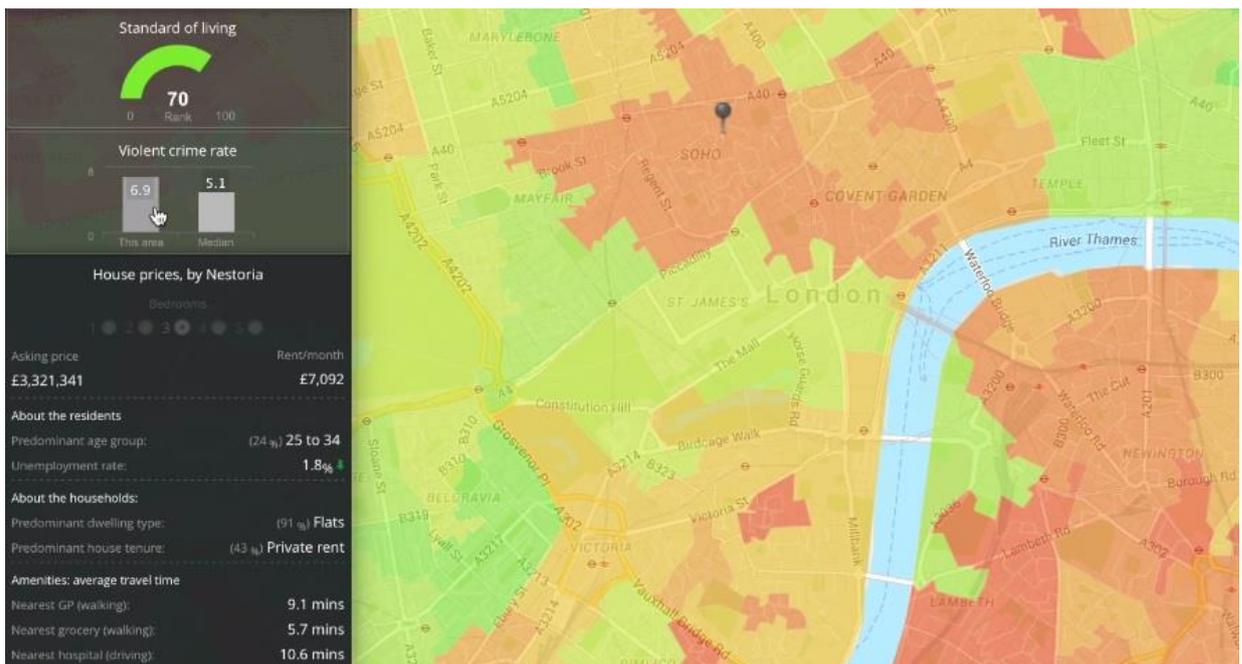
17 Resultados elecciones generales en el municipio de Madrid; fuente: El Confidencial.LAB



18 Mapa de representación de las multas de tráfico en Madrid dentro de la M30; fuente: ElPaís.com

Por otro lado, además del software *ArcGis* y de las múltiples tablas de datos de *Excel* y los análisis realizados con el software de estadística *SPSS*, para realizar los análisis de regresión y realizar el Análisis Exploratorio de Datos Espaciales (AEDE) hemos utilizado el software libre *GeoDa*, un paquete de software gratuito que lleva a cabo el análisis espacial de datos, geovisualización, autocorrelación espacial (*Moran* y *LISA*) y modelado espacial.

Otro de los programas informáticos de análisis espacial es *SaTScan*, un software gratuito que analiza datos espaciales, temporales y de espacio-tiempo utilizando datos de estadísticas, y que está diseñado para, entre otros propósitos, la realización de la vigilancia geográfica de enfermedades y para detectar agrupaciones espaciales o espacio-temporales y ver si son estadísticamente significativas, por ejemplo probar si una enfermedad está distribuida al azar en el espacio, en el tiempo o en el espacio y en el tiempo a la misma vez, además de para realizar vigilancia de enfermedades de forma periódica para la detección temprana de brotes (vigilancia de la salud). Este software también puede ser utilizado para problemas similares en otros campos distintos de la salud como la arqueología, la astronomía, la botánica, la criminología, la ecología, la economía, la ingeniería, la silvicultura, la genética, la geografía, la geología, la historia, la neurología o la zoología.

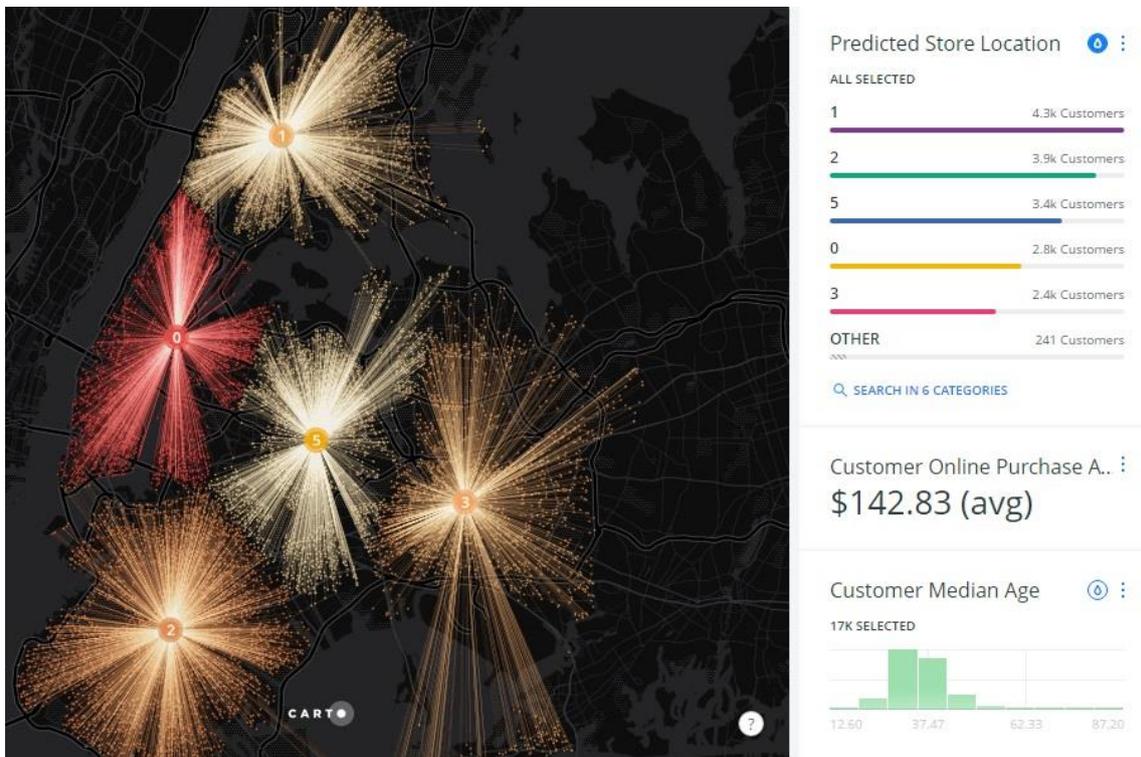


19 Análisis de datos espaciales de la ciudad de Londres; fuente: *OpenStreetMap.org*

Además de los SIG comentados y del archiconocido **Google Maps**, que también puede utilizarse como una sencilla herramienta de georreferenciación, otros proyectos interesantes han surgido en los últimos años en este campo.

**OpenStreetMap** ([www.openstreetmap.org](http://www.openstreetmap.org)) por ejemplo, es un proyecto colaborativo para crear mapas libres que se crean utilizando información geográfica capturada con dispositivos GPS móviles, ortofotografías (similares a las tomadas por Google) y otras fuentes libres. Para hacerse una idea del ámbito colaborativo y de la extensión de los SIG en los últimos tiempos, en octubre de 2014 estaban registrados en torno a 1.840.000 usuarios.

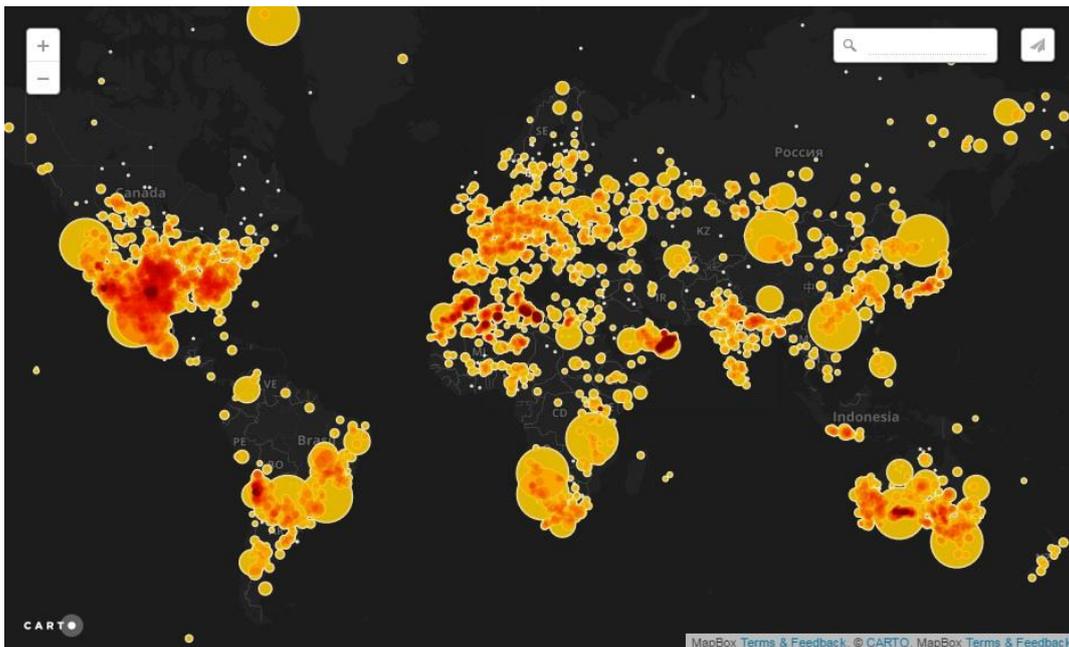
Los usuarios registrados en **OpenStreetMap** pueden subir sus trazas desde el GPS y crear y corregir datos vectoriales mediante herramientas de edición creadas por la comunidad. En la mayoría de los países la información geográfica pública no es de libre uso, sino que hay que registrarse para el acceso a la misma. En cualquier caso, el alcance de estas herramientas que funcionan con la filosofía de conocimiento colaborativo, está creciendo de manera exponencial en los últimos años.



20 Mapa sobre potenciales compradores en zonas comerciales; fuente: web de CartoDB.

Por otro lado, en los últimos años también han surgido iniciativas comerciales como **MapShare** de **TomTom** o **Map Maker** de **Google**, orientadas a animar a los usuarios de sus servicios a completar los mapas generados por su aplicación, actualizando y corrigiendo su cartografía y agregando nuevos datos. En la mayoría de estos casos los usuarios no tienen derecho alguno sobre esa cartografía o datos que están añadiendo o editando, pasando a ser sus contribuciones propiedad de dichas empresas.

Por último, otra potente herramienta en este ámbito es **CartoDB** ([www.carto.com](http://www.carto.com)), una iniciativa española y estadounidense (sus sedes están en Madrid y en Nueva York) creada para que trabajar con datos geoespaciales y grandes cantidades de datos (*Big Data*) sea más rápido, sencillo y más accesible. La compañía busca ser una plataforma de *inteligencia de localización* gracias a las herramientas que ofrece para el análisis y visualización de datos que no requieren de gran experiencia en el manejo de los SIG. Así, en vídeos como el del siguiente enlace (<https://vimeo.com/59791629>) se observa cómo un periodista del *New York Times* ha elaborado un mapa dinámico de los meteoritos caídos en la Tierra a través de diversas fuentes de datos, algo muy común en medios de comunicación escritos que últimamente muestran casi a diario mapas con múltiples datos e informaciones para ilustrar mejor las noticias.



*21 Georreferenciación de los meteoritos caídos en la Tierra y su tamaño; fuente: CartoDB.*

Desde el editor de *CartoDB*, así como en otros softwares similares, se pueden importar varios tipos de datos:

.CSV - valores separados por comas -

.SHP - ESRI shapefiles -

.KML, .KMZ - formatos de *Google Earth* -

.XLS, .XLSX - formatos de *Microsoft Excel* -

.GEOJSON - *GeoJSON* -

.OSM - *OpenStreetMap* -

.ODS - Hoja de cálculo de *OpenDocument* y *OpenOffice* -

El concepto de **Location Intelligence** que ofrece *CartoDB* persigue utilizar la información gráfica y espacial para analizar entornos socioeconómicos y estudiar dependencias y relaciones espacio-temporales y así avanzar el estado del conocimiento en el campo de, entre otros, la econometría espacial. Así, podemos observar numerosos estudios realizados para analizar determinadas zonas geográficas o urbanas en busca de la mejor localización para establecer un comercio en base a numerosas variables, o identificar la mejor ruta a seguir en tiempo real analizando aspectos del día a día como los atascos de tráfico o las obras de construcción.

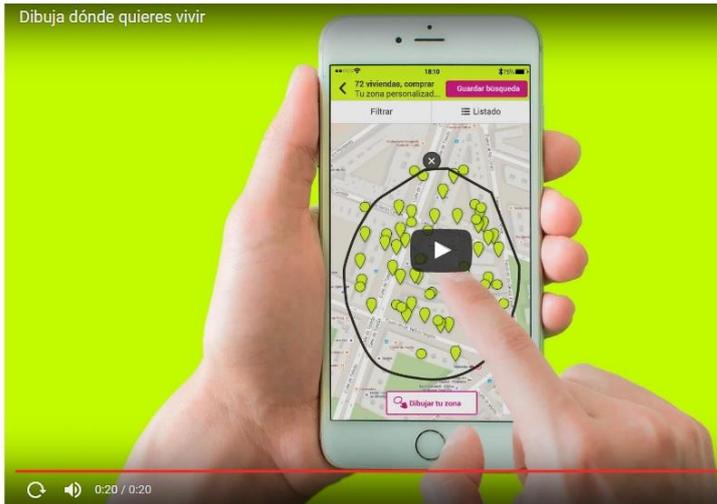
*Localización de inteligencia* es un término poco utilizado todavía en español y lo que significa y cómo puede beneficiar a las empresas, instituciones e individuos puede no ser demasiado evidente ni fácil de explicar, pero en definitiva es algo más que un análisis espacial mediante SIG, es la capacidad de visualizar datos espaciales para identificar patrones y analizar las relaciones entre los diversos datos y variables que se puedan georreferenciar de algún modo. Algunas empresas del mundo anglosajón (finanzas, bienes raíces, logística operacional) ya utilizan estas tecnologías para obtener una ventaja competitiva en el mercado.

## USING DATA ANALYSIS TO PROVIDE INSIGHTS INTO RENT RATES

CARTO visualized with detailed U.S. Census data, from CARTO's Data Observatory, city Airbnb prices nationwide. Then, compared median Airbnb prices to median rent for every block in New York City. We find that clusters of high profitability in Manhattan center around Chinatown, East Harlem, and Manhattan Valley just west of Central Park. Williamsburg was the only large high-profitability cluster in Brooklyn, with a median Airbnb listing able to pay median rent in just five days.



22 Mapa sobre precios de alquileres de viviendas de Airbnb; fuente: web de CartoDB.

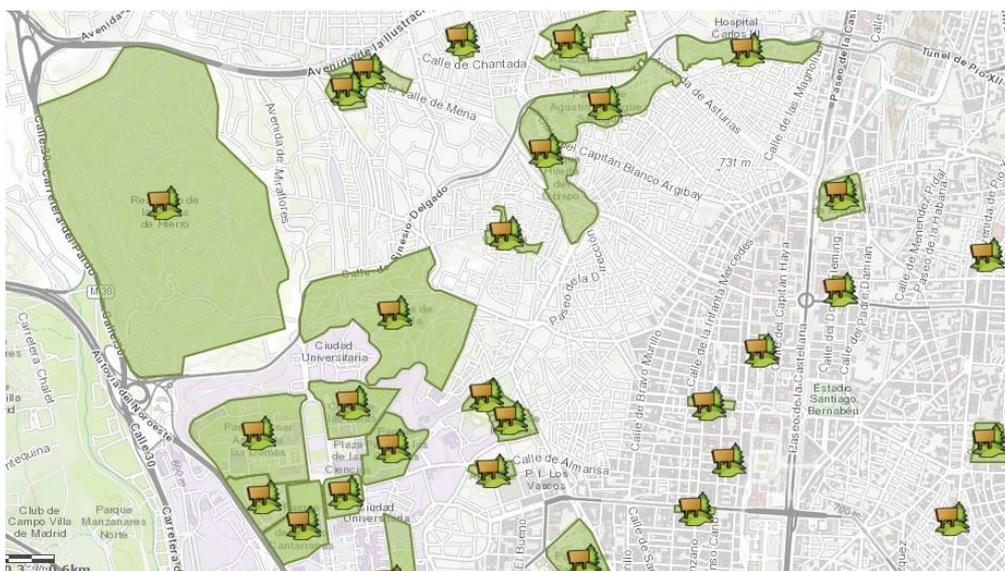


### Dibuja dónde quieres vivir, la novedosa funcionalidad de idealista para buscar casa

23 Aplicación para Smartphone de búsqueda de viviendas mediante georreferenciación; fuente: Idealista.com.

En la web [www.carto.com/solutions/real-estate](http://www.carto.com/solutions/real-estate) tenemos una aplicación especializada en el sector inmobiliario en la que se pueden analizar múltiples bases de datos relacionados con el urbanismo y el sector de la vivienda así como para visualizar las ubicaciones de propiedades comerciales y residenciales de cualquier área urbana utilizando indicadores de datos detallados como la densidad de población y la producción económica.

De este modo, junto a los modelos de regresión utilizados en el trabajo, hemos utilizado los SIG para analizar la estructura espacial de la ciudad de Madrid a través de mapas y georreferenciando diversos elementos para analizar variables de carácter espacial y estudiar relaciones de dependencia y correlación entre ellas.



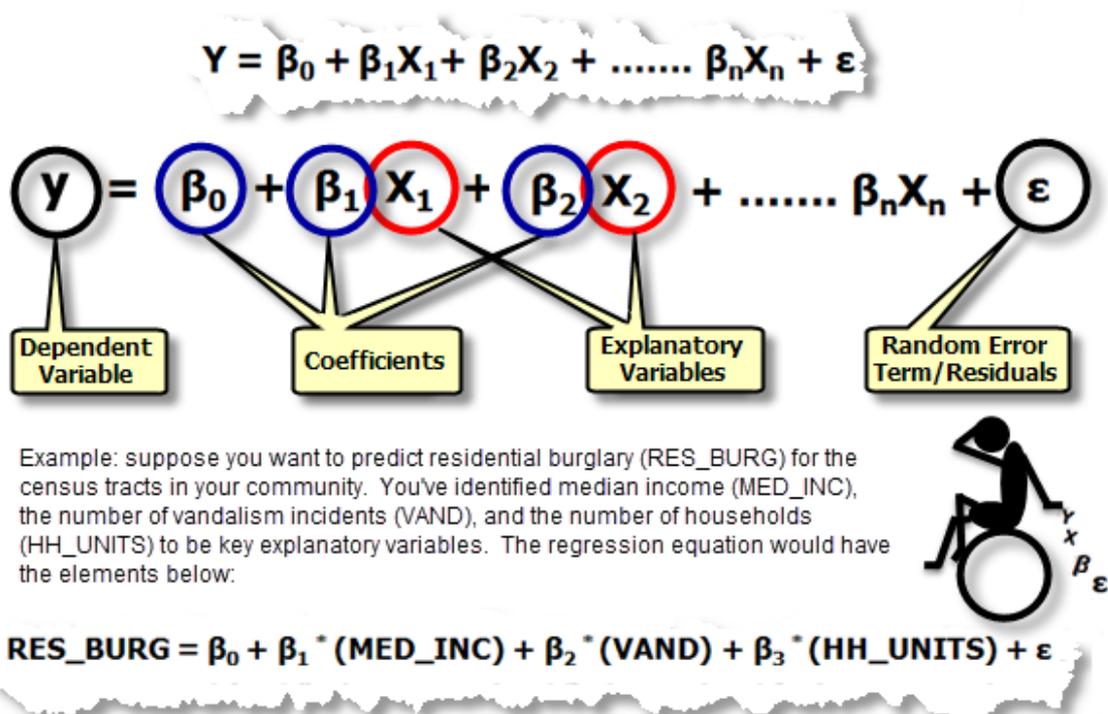
24 Georreferenciación de zonas verdes en la zona analizada del municipio de Madrid; fuente: elaboración propia.

## Las regresiones hedónicas y espaciales

En este apartado se pretende explicar los métodos utilizados en el análisis de los datos de este TFM mediante regresiones, analizando los distintos tipos de regresiones para el análisis estadístico y el tratamiento de datos espaciales. Para la elaboración de este apartado se ha utilizado la información de varios tutoriales sobre econometría espacial y análisis de regresión disponibles en la web de ARCGIS online, el sistema de información geográfica utilizado en el análisis espacial de este trabajo de investigación.

Para comenzar, explicaremos en qué consiste un análisis de regresión, que no es otra cosa que una fórmula matemática que se aplica a las variables explicativas para prever mejor la variable dependiente que intenta modelar. En el caso de este TFM la variable dependiente es siempre el precio de la vivienda y las explicativas son las diversas características de las viviendas, así como las distancias calculadas.

A través del análisis y el cálculo estadístico, cada variable independiente (las X) está asociada a un coeficiente de regresión que describe la longitud y el signo de la relación de esa variable con la variable dependiente (la Y). Una ecuación de regresión puede tener un aspecto como el de la imagen que acompaña a esta información (Y es la variable dependiente, las X son las variables explicativas y las  $\beta$  son coeficientes de regresión). Las  $\beta$  son los resultados del análisis y lo que nos da información acerca del peso que cada variable explicativa tiene sobre la variable dependiente.



*25 Elementos de una ecuación de regresión de OLS; fuente: tutorial de regresiones de ArcGIS*

En economía la técnica de regresión más conocida es la de *Mínimos Cuadrados Ordinarios* (OLS por sus siglas en inglés). Además, vamos a explicar el concepto de *Regresión Hedónica* y el de *Regresión Geográficamente Ponderada* (GWR por sus siglas en inglés).

La regresión hedónica es un método para estimar la preferencia revelada basado en la descomposición de un bien económico (en este caso la vivienda) en sus características más importantes y en el análisis de la contribución al valor agregado de cada una de tales características. De este modo, se descompone el artículo que es investigado en sus características constitutivas (tamaño de la vivienda, número de baños, habitaciones...), y se obtienen estimaciones del valor contributivo de cada característica (los coeficientes  $\beta$  de la regresión).

Los modelos hedónicos se utilizan generalmente en la evaluación de bienes raíces y también en la estimación de los costos relacionados con la calidad general del medio ambiente en términos de contaminación del aire, contaminación del agua, ruido, instalaciones ambientales que incluyen lugares estéticos y cercanía a los sitios de recreación como parques, playas, etc.

Las primeras aplicaciones del método de los precios hedónicos datan de los años veinte del siglo pasado en Estados Unidos. Se utilizó originalmente para estimar el valor de propiedades agrarias en función de la fertilidad del suelo, la distancia a la ciudad del mercado, el valor de la construcción, etc. También en sus inicios tuvo aplicaciones en el fructífero sector del automóvil.

El método de precios hedónicos refleja el *bienestar* o *disfrute* (que es lo que significa originalmente hedónico en griego) que ofrecía cada componente o característica del bien analizado. Así, cuando los precios de un determinado bien suben porque las prestaciones del bien son mejores, esta inflación se conoce como inflación hedónica.

En los años sesenta y setenta el método tomo otro impulso con las contribuciones más formales de otros economistas, en ese momento ya con acceso a ordenadores con nueva capacidad de cálculo. El desarrollo del análisis de regresión contribuyó lógicamente en gran medida al desarrollo del método. Actualmente, con el desarrollo de nuevos y potentes softwares de análisis espacial, se están incorporando cada vez más los factores espaciales al análisis de regresiones hedónicas, como se ha pretendido en este TFM.

Según los apuntes del profesor Benjamín López Ortiz, de la Facultad de Economía de la Universidad Nacional Autónoma de México (UNAM), que explica en sus clases las regresiones hedónicas, un ejemplo ilustrativo para entender el método de precios hedónicos sería el siguiente:

Supongamos que estamos en proceso de alquilar una vivienda, digamos un piso de tres habitaciones en una ciudad cualquiera. De los apartamentos que visitamos, dudamos entre dos que tienen características muy similares, pero difieren en que uno se halla en una calle especialmente ruidosa, mientras que el otro está un par de calles más allá, con mucho menos ruido.

Supongamos que la primera vivienda soporta una media de 60 decibelios durante el día y la segunda 45. El precio del primero es de 600€ de alquiler mensual y el segundo 650€.

Si las únicas diferencias relevantes entre los dos pisos son el precio y los decibelios, y nos decidimos por alquilar el primero de los dos, ¿Qué estamos revelando de nuestras preferencias sobre el ruido?

Pues que soportar durante el día en casa una diferencia de 60 a 45 decibelios no nos molesta lo suficientemente como para pagar 50€ más al mes. Nuestro valor por esa diferencia de ruido es inferior a esos 50€ mensuales. Si hubiéramos optado por el otro piso, sería porque nuestro valor por el beneficio de ahorrarnos la diferencia de decibelios sería de 50 o más euros mensuales. Así, las preferencias que revelamos con nuestra decisión de alquiler en el mercado de las viviendas dan pistas sobre los valores que para nosotros tienen unos bienes (en este ejemplo, la ausencia del ruido) para los que no tenemos mercados en los que observar directamente esas preferencias.

Desafortunadamente, no es frecuente hallar situaciones donde dos bienes varíen para las preferencias de las personas sólo en la característica cuyo valor nos interesa valorar. De hecho, cuando visitamos distintas viviendas para alquilar o comprar, suelen variar entre ellas en bastantes características a la vez, de ahí que el método sea mucho más complejo que en este sencillo ejemplo. En trabajos de investigación con múltiples variables explicativas, como es el caso de este TFM, se analizan pormenorizadamente variables de las viviendas que nos ayudan a explicar la preferencia revelada del consumidor a la hora de comprar una vivienda en la ciudad de Madrid, teniendo además en cuenta los factores espaciales, lo que dificulta los modelos y hace aún más complejo los análisis de regresión.

Precisamente por esta complejidad a la hora de tomar en cuenta factores de índole espacial, se hacen necesarias las denominadas regresiones espaciales, denominadas también regresiones geográficamente ponderadas (GWR por sus siglas en inglés).

Los datos espaciales exhiben propiedades que hacen difícil cumplir con las presuposiciones y los requisitos de los métodos estadísticos tradicionales, como es la regresión de Mínimos Cuadrados Ordinarios (OLS). En general, las entidades geográficas están autocorrelacionadas espacialmente, esto significa que las entidades cercanas tienden a ser más similares que las entidades que están más lejos (*Ley de Tobler*). Esto es lo que se denomina autocorrelación espacial, y hay que tenerlo en cuenta a la hora de analizar elementos cuya ubicación geográfica pueda influir o determinar especialmente la variable o fenómeno analizado.

*Los modelos estadísticos tradicionales pueden mostrar problemas de especificación cuando no consideran la presencia de autocorrelación espacial en los datos. Una consecuencia importante es el riesgo de estimar coeficientes ineficientes para representar la magnitud de la relación entre las variables, por lo que las pruebas de significación estadística sobre los mismos serán cuestionables debido a la inflación en los errores estándar. Por esta razón, antes de cualquier interpretación de los resultados de un análisis regresión, es necesario someter los residuos a test específicos (como el I de Moran o la C de Geary) para demostrar estadísticamente su aleatoriedad espacial. En caso contrario, es decir, en presencia de*

*autocorrelación espacial, es recomendable evaluar el empleo de alguna técnica diseñada específicamente para hacer frente a este tipo de problemas, como la Regresión Geográficamente Ponderada (GWR). (Gutiérrez Puebla, 2012)*

Fueron *Brunsdon et al. (1996)* los que introdujeron este término para aludir a una **familia de modelos de regresión “ajustados al espacio”, donde fuera posible observar las variaciones espaciales de los parámetros estimados y con ello saber dónde y cuánto es el efecto de una variable explicativa sobre la dependiente.** Se trata de ajustar tantas regresiones como observaciones (unidades espaciales) se consideren en el análisis, en base al concepto de *distance decay* (se da más peso a las observaciones más próximas y menos a las más lejanas).

En los modelos utilizados en los modelos de regresión de este TFM se han tenido en cuenta variables espaciales mediante una *matriz de pesos* en las variables *weight\_4* y *w\_size* que incorporan el efecto de la influencia espacial de las 4 viviendas más cercanas a cada observación en el precio y en el tamaño de las viviendas de la zona analizada.

## Autocorrelación espacial global y local

Este apartado pretende explicar las técnicas estadísticas que utilizaremos para analizar la presencia de patrones espaciales en la formación del precio de la vivienda en la zona analizada.

La metodología que hemos utilizado en este TFM está basada en el denominado como *Análisis Exploratorio de Datos Espaciales* (en adelante AEDE), que consiste en una serie de técnicas para visualizar y calcular la autocorrelación espacial, si es que la hay. El AEDE se puede definir como el conjunto de técnicas que describen y visualizan las distribuciones espaciales, identifica ubicaciones atípicas (*outliers*) e identifica patrones de asociación según la localización como los conglomerados (*clusters*) o los puntos calientes (*hot spots*). En definitiva, se trata de un conjunto de técnicas que muestran estructuras espaciales y cualquier forma de heterogeneidad geográfica.

Parece haber una gran diferencia entre cómo ve un estadístico tradicional el fenómeno de la autocorrelación espacial y cómo lo ve un estadístico espacial (un analista de SIG). El estadístico tradicional la ve como algo malo que se debe eliminar de los datos (a través del remuestreo, por ejemplo) porque la autocorrelación espacial viola las presuposiciones subyacentes de varios métodos estadísticos (no espaciales) tradicionales. Sin embargo, para el geógrafo o analista de SIG, la autocorrelación espacial es evidencia de procesos espaciales subyacentes importantes en el trabajo; es un componente integral de los datos.

### Técnicas de Análisis Exploratorio de Datos Espaciales

Perspectiva Econometría Espacial	
<b>Visualización de distribuciones espaciales</b>	<ul style="list-style-type: none"><li>• <i>Box map</i>,</li><li>• Histograma,</li><li>• Análisis de la varianza exploratorio espacial</li></ul>
<b>Visualización de asociación espacial global</b>	<ul style="list-style-type: none"><li>• Gráficos del retardo espacial,</li><li>• Mapa y <i>Scatterplot</i> de Moran</li></ul>
<b>Visualización de asociación espacial local</b>	<ul style="list-style-type: none"><li>• Mapas LISA,</li><li>• <i>Outliers</i> en el <i>Scatterplot</i> de Moran</li></ul>
<b>Asociación espacial multivariante</b>	<ul style="list-style-type: none"><li>• <i>Scatterplot</i> multivariante de Moran</li></ul>
<b>Heterogeneidad Espacial</b>	<ul style="list-style-type: none"><li>• Mapa,</li><li>• Histograma de Frecuencias,</li><li>• Diagrama de Dispersión</li></ul>

Fuente: (Anselin, 1988)

26 Técnicas empleadas en el AEDE; fuente: Luc Anselin (1988)

Los procesos espaciales y las relaciones espaciales evidentes en los datos son de un gran interés y es una de las razones por las que los usuarios de SIG se interesan en el análisis de datos espaciales. Sin embargo, para evitar un tipo de influencia que cuente más en su modelo, debe identificar el conjunto completo de variables explicativas que efectivamente capturen la estructura espacial inherente en su variable dependiente. Por ello, tanto en la utilización de los modelos de regresión con factores espaciales como en el uso de los SIG es muy importante definir correctamente las variables y redefinir el modelo tantas veces como haga falta, teniendo en cuenta las características específicas de las variables espaciales.

Para comprobar si se da el fenómeno de la autocorrelación espacial global se pueden utilizar diferentes índices, pero el más conocido y utilizado es el **Índice de Moran**. Es un factor que se calcula para **detectar la correlación entre los valores de una variable en una localización y los valores en su entorno**. Puede ser utilizado para contrastar la hipótesis nula de ausencia de autocorrelación espacial (la localización espacial de los valores de la variable de forma aleatoria) frente a la hipótesis alternativa de existencia de autocorrelación espacial (positiva o negativa) (Upton y Fingleton, 1985).

Así, el índice mide la autocorrelación espacial basada en las ubicaciones y los valores de las entidades simultáneamente. Dado un conjunto de entidades y un atributo asociado (en nuestro caso son las viviendas y el atributo es el precio), **evalúa si el patrón expresado está agrupado, disperso o es aleatorio**. Se calcula el valor del Índice I de Moran (IM) así como una puntuación z y un valor P para evaluar la significancia de ese índice.

Si los valores del conjunto de datos (*dataset*) tienden a agruparse espacialmente (los valores altos se agrupan cerca de otros valores altos; los valores bajos se agrupan cerca de otros valores bajos), el Índice de Moran será positivo.

Cuando los valores altos rechazan otros valores altos y tienden a estar cerca de valores bajos, el índice será negativo. Si los valores positivos de los productos cruzados equilibran los valores negativos de los productos cruzados, el índice será cercano a cero. El numerador está normalizado por la varianza de modo que los valores del índice oscilan entre -1 y 1.

En estadística, el índice desarrollado por Patrick Moran explica que la autocorrelación espacial es más compleja que una dimensión de autocorrelación simple ya que la correlación espacial es multidimensional (considera las dos o tres dimensiones del espacio) y, además, es multidireccional. La fórmula del índice de Moran es la siguiente:

$$I = \frac{N}{\sum_i \sum_j \omega_{ij}} \frac{\sum_i \sum_j \omega_{ij} (X_i - \bar{X})(X_j - \bar{X})}{\sum_i (X_i - \bar{X})^2}$$

donde,

$N$  es el número de unidades espaciales indexados por  $i$  y  $j$

$X$  es la variable de interés

$\bar{X}$  es la media de  $X$

$\omega_{ij}$  es un elemento de una matriz de pesos espaciales

Otro de los estadísticos más utilizados en el análisis de la dependencia espacial es la denominada *C de Geary*, pero mientras que el contraste del *I de Moran* es una medida de autocorrelación espacial global la *C de Geary* es más sensible a la autocorrelación espacial local.

En lo que se refiere a la autocorrelación espacial local, el indicador de correlación espacial más utilizado y el utilizado en esta investigación es el conocido como **Local Indicator of Spatial Association LISA** (Anselin, 1995), que indica la semejanza del valor observado en cada localización con los de su entorno. Así, si el valor de este estadístico para una localización es positivo (y significativo), la unidad geográfica puede ser considerada como el centro de un *clúster* (grupo o conjunto) bien por ser una zona de valores elevados rodeada de valores elevados (*high-high*) o bien por ser una zona de valores bajos rodeada de zonas de valores bajos (*low-low*). Por el contrario, si el valor es negativo la unidad espacial puede considerarse como una zona atípica espacialmente, bien por ser una zona de alta (baja) incidencia rodeada de zonas de baja (alta) incidencia: (*high-low*) o (*low-high*). En el caso de este TFM, a pesar de poder realizarse con diferentes softwares, el AEDE se ha realizado con *GeoDA*, y los resultados de IM y LISA se encuentran más adelante junto con los mapas que los reflejan gráficamente (*boxplot* y *scatterplot*).

## DATOS

Para la realización de este trabajo se han obtenido numerosos datos de muy diversas fuentes, hasta llegar a construir una amplia base de datos que incluye una extensa información muestral de 4.100 observaciones de precios de viviendas obtenidas de la página web *idealista.com*.

Además, para la obtención de datos georreferenciados de la ciudad de Madrid hemos acudido al portal de “Datos Abiertos” del Ayuntamiento de Madrid, una web dedicada a promover el acceso a los datos del gobierno municipal e impulsar el desarrollo de herramientas creativas para atraer y servir a los ciudadanos de la ciudad de Madrid.

A continuación se explica el proceso seguido para la obtención y construcción de la base de datos empleada en el trabajo así como información acerca de las diversas fuentes de donde extraer información y datos sobre el precio de la vivienda en España.

### Fuentes de datos sobre precios de la vivienda

Existen diversas fuentes y bases de datos para extraer información sobre precios de vivienda a las que podemos acudir. Las páginas web *precioviviendas.com* y *preciosdevivienda.es* son dos ejemplos de donde se puede extraer información útil y actualizada sobre precios de viviendas e inmuebles en general.

La web *precioviviendas.com* es una web conformada por profesionales del sector de la valoración de inmuebles y las tasaciones cuyo sistema de valoración es certificado periódicamente por el Departamento de Estadística e Investigación Operativa de la Facultad de Ciencias Matemáticas de la Universidad Complutense de Madrid, cuyos investigadores han desarrollado una metodología de valoración basada en regresiones hedónicas que tienen en cuenta las características propias del inmueble para proceder a su valoración, algo parecido al estudio realizado en este trabajo fin de máster.

Esta web trabaja con los datos obtenidos en la web oficial del catastro (*www.sedecatastro.gob.es*) y también con la base de datos disponible en la división de secciones censales del Instituto Nacional de Estadística (*www.ine.es*). Para ello realizan procesos de *Data Cleaning* y de *Data Mining* continuos, en los que obtienen una base de datos completa y actualizada de los municipios en los que ofrecen sus servicios (actualmente sólo ofrecen su servicio en las ciudades de Madrid, Barcelona, Sevilla y Valencia).

En la búsqueda de datos sobre precios de la vivienda cabe recordar que existen varios valores a la hora de establecer precios de un mismo inmueble, así, por ejemplo, en la web del catastro sólo podremos conseguir datos referentes a la valoración catastral de las viviendas, mientras que para obtener precios o valores de mercado podremos acceder a las páginas de empresas inmobiliarias como **Fotocasa** o **Idealista** (la web de donde se ha obtenido la base de datos para este trabajo de investigación).

Lo que ofrecen las páginas web como *precioviviendas.com* es la posibilidad de acceder a una valoración de los inmuebles mediante metodologías contrastadas que recogen los datos de diversas fuentes y que obtienen, tras el análisis y el estudio del inmueble (situación, características...) una valoración aproximada del valor de mercado del mismo. En el caso de esta web, la metodología utilizada viene avalada por los miembros del grupo de investigación de la Universidad Complutense "Modelos estadísticos para la valoración de activos inmobiliarios", cuya colaboración con la empresa Aplicaciones Estadísticas y Consultoría, S.L. ha dado lugar a la creación de un *Modelo de Valoración Automatizada (AVM)* para viviendas ubicadas en todo el territorio nacional que incluye un *Sistema de Medición de la Fiabilidad (SMF)* que mide de manera semestral y anual la fiabilidad de los valores automáticos obtenidos por dicho modelo en contraste con los valores de mercado obtenidos en valoraciones con visita física a las viviendas efectuadas por la compañía **Sociedad de Tasación**, que es una de las grandes compañías españolas que opera en el campo de las valoraciones de activos inmobiliarios ([www.st-tasacion.es](http://www.st-tasacion.es)).

Cabe destacar que, según se recoge en la propia web, la herramienta calcula el valor de mercado de las viviendas sin realizar visita física o análisis de documentación alguna, por lo que sus resultados carecen de toda validez legal y no se recomienda el uso de estos precios de viviendas en caso de operaciones económicas sobre las mismas, recomendándose la obtención de una valoración por tasadora independiente. En la siguiente imagen podemos observar un ejemplo de un informe realizado por la web *precioviviendas.com* acerca de la valoración de un inmueble de la ciudad de Madrid, donde obtenemos información detallada del inmueble así como datos acerca de la evolución de los precios en la zona y otros datos de interés como el precio/m<sup>2</sup>.

**precio**  
**viviendas**  
www.precioviviendas.com

Informe N° 5.414

**Dirección del inmueble a valorar:** CL CRISTO DE LA LUZ 1, MADRID, MADRID

**Fecha de la valoración:** 27 de mayo de 2016

**DATOS UTILIZADOS EN LA VALORACIÓN**

Tipología:	Piso	Superficie parcela:	-	Reforma:	No
Altura:	Planta 3	Superficie adoptada:	51 m <sup>2</sup>	Año reforma:	-
Vistas:	Normales	Nº de baños/aseos:	1	Intens. Reforma:	-
Ascensor:	Sí	Antigüedad:	55 años		
Dúplex:	No	Calidad construcción:	Media		
Ático con terraza:	No	Calidad Ubicación:	Igual		

**Datos catastrales:**

Referencia:	0605501VK4700F0011FA	Uso:	Residencial	Antigüedad:	1960
Superficie total:	51 m <sup>2</sup>	Superficie construida:	47 m <sup>2</sup>	Superficie parcela:	209 m <sup>2</sup>

**VALORACIÓN**

Superficie construida por ppz: 51 m<sup>2</sup>

Valor del inmueble: 67.269 €

Variación de Confianza: 1.253 €/m<sup>2</sup> - 1.385 €/m<sup>2</sup>

**COMENTARIO**

Una vivienda con similares características a la valorada situada en su misma zona tiene un valor aproximado de 1.319 €/m<sup>2</sup>.  
El Grado de Confianza en dicho valor es Alto.

La fiabilidad del sistema de valoración ha sido certificada por el Departamento de Estadística e Investigación Operativa de la Facultad de Matemáticas de la Universidad Complutense de Madrid.

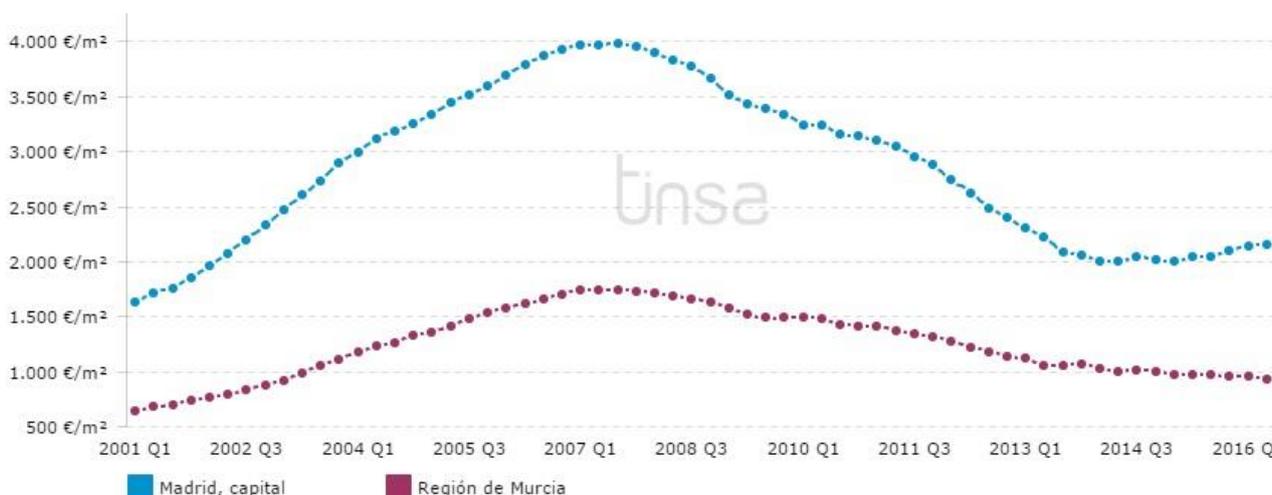
27 Informe con datos sobre valoración de una vivienda en Madrid; fuente: web de *precioviviendas.com*.

Por otro lado, las páginas web de empresas dedicadas a la tasación también ofrecen sus servicios y en ellas se puede obtener información acerca de valoraciones de mercado de bienes inmuebles. Además de la empresa *Sociedad de Tasación*, quizá la mayor empresa que opera en España en el campo de las valoraciones inmobiliarias sea **TINSA**, una compañía multinacional dedicada a la valoración y al asesoramiento inmobiliario en general que cuenta entre sus clientes con entidades financieras de los países donde opera, empresas de múltiples sectores, Administraciones Públicas y particulares.

*TINSA* dispone, además, de un servicio de tasación online mediante la herramienta *STIMA* ([www.tinsa.es/store/tasacion-online-stima](http://www.tinsa.es/store/tasacion-online-stima)), que a diferencia de otras herramientas online disponibles, que utilizan datos procedentes de oferta de compraventa de viviendas, emplea datos reales de tasaciones elaboradas por los técnicos y profesionales de la compañía, por lo que la valoración se ajusta de forma más veraz al valor real de mercado del inmueble.

Además de los servicios de tasación ofrecidos por sus técnicos y el servicio online, *TINSA* elabora el índice *Tinsa IMIE (Informe de los Mercados Inmobiliarios Españoles)*, que es una buena herramienta para analizar la evolución del valor de la vivienda en el mercado español. Este informe se elabora mensualmente y analiza la base de datos de tasaciones de la compañía para calcular el precio medio de la vivienda terminada en las Comunidades Autónomas, provincias y capitales de España. Por ello, proporciona información útil para realizar análisis de inversiones, realizar operaciones de rentas inmobiliarias, identificar la tendencia del mercado inmobiliario en una determinada zona (desagrega los datos por Comunidades Autónomas, provincias y capitales de provincia) e incluso comparar la evolución del mercado residencial con la de otros sectores, ya que ofrece numerosos índices y ratios económicos relativos al sector inmobiliario.

En estas imágenes se puede ver algunas de las gráficas del *IMIE* referentes al informe de Julio de 2016.



**28 Evolución del precio por metro cuadrado en la ciudad de Madrid y en la Región de Murcia; fuente: web de TINSA.**

SEGUNDO TRIMESTRE 2016	
	Var. interanual
Media nacional	0,8% ↑
Comunidad de Madrid	5,4% ↑
<b>Resumen Madrid, capital</b>	
Variación interanual	5,1% ↑
Diferencia desde máximos	46,0% ↓
Var. Acum. Q2 2016*	2,8% ↑
Var. Acum. Q2 2015*	1,5% ↑
Tiempo medio de venta	5,2 meses
Precio medio	2.147 €/m <sup>2</sup>

29 Datos sobre el precio de las viviendas en la ciudad de Madrid; fuente: web de TINSA.

La página web preciosdevivienda.es es un proyecto que elabora el denominado *Panel de Precios de la Vivienda*, y surge, según la descripción ofrecida en el portal, “de la necesidad de presentar en un espacio común la gran variedad de información que sobre precios de la vivienda se difunde periódicamente en España”.

El Panel recopila la información que elaboran y difunden diferentes instituciones, públicas y privadas, así como asociaciones vinculadas al sector inmobiliario. La información ha sido ordenada de forma sencilla, para que cualquier usuario pueda conocer y comparar, además de los datos

producidos por cada una de las instituciones, las principales características metodológicas de las respectivas estadísticas, su periodicidad y el ámbito territorial de aplicación. De este modo, esta herramienta permite conocer de manera fiable los precios de la vivienda a nivel nacional elaborados con la base de datos de las principales instituciones que estudian y analizan el ámbito inmobiliario.

Estas instituciones son, por el ámbito público el **Ministerio de Fomento** ([www.fomento.gob.es](http://www.fomento.gob.es)) y el **Instituto Nacional de Estadística** ([www.ine.es](http://www.ine.es)) y las instituciones de ámbito privado *Idealista.com* (el portal inmobiliario del que se ha extraído la base de datos de este TFM), las empresas de valoración inmobiliaria *Sociedad de Tasación* y *TINSA* (de las que hemos hablado anteriormente), el **Consejo General del Notariado** ([www.notariado.org](http://www.notariado.org)) y el **Colegio de Registradores** ([www.registradores.org](http://www.registradores.org)).

El Instituto Nacional de Estadística (INE) elabora el índice de precios de vivienda (IPV), un índice que tiene como objetivo medir la evolución de los precios de compraventa de las viviendas de precio libre a lo largo del tiempo. El IPV responde a la demanda de información en el ámbito de la producción de estadísticas armonizadas de la Unión Europea, por lo que uno de sus objetivos es servir de elemento de comparación entre los estados miembros de la UE en lo que se refiere a la evolución de los precios de vivienda. El IPV elaborado por el INE utiliza la información sobre escrituraciones procedente de las bases de datos proporcionadas por el Consejo General del Notariado. El IPV, por tanto, se alimenta de los valores y superficies declarados en las escrituras formalizadas ante notario recogiendo el precio de mercado del metro cuadrado de la vivienda. Cabe destacar que éste y otros índices presentan un problema de fiabilidad en cuanto a que la componente del mercado negro en el sector inmobiliario no es nada desdeñable, y ya que los datos de la mayoría de series estadísticas provienen de registros oficiales, los índices relativos a precios de mercado pueden no estar rigurosamente acorde a los valores de mercado reales del sector.

Como vemos, las fuentes estadísticas sobre precios de la vivienda son principalmente las elaboradas por el INE, el Consejo General del Notariado y el Colegio de Registradores, ya que la información elaborada por otros portales web y por las empresas de tasación bebe de la información recogida y analizada por estas instituciones. Dependiendo de la metodología utilizada y el fin que se le quiera dar a estos índices y bases de datos estas fuentes varían ligeramente. Lo que sí es cierto es que, como alertan los investigadores José María Montero y Gema Fernández en su artículo "*La importancia de los efectos espaciales en la predicción del precio de la vivienda. Una aplicación geoestadística en España*", ninguna de estas fuentes estadísticas incorpora efectos espaciales en la elaboración de los índices, ya que la mayoría de ellos se generan a partir de datos procedentes de las transacciones realizadas y de los datos provenientes de los diferentes registros. Así, las estadísticas sobre precios de vivienda en España requieren mejoras en su metodología y en el modo en el que consideran las diferentes variables a tener en cuenta a la hora de calcular los índices de precios. Entre otras cuestiones, la consideración de los efectos espaciales (como los tenidos en cuenta en la elaboración de los modelos de este trabajo de investigación).

El Ministerio de Fomento elabora trimestralmente el precio medio de la vivienda tasada; el Instituto Nacional de Estadística (INE) elabora, como hemos indicado anteriormente, el índice de precios de vivienda (IPV); el Colegio de Registradores de la Propiedad pública anualmente el Índice de Precios de la Vivienda de Ventas Repetidas (IPVVR); el Consejo General del Notariado elabora series mensuales del precio de la vivienda libre y de protección oficial. Por otro lado, algunas sociedades tasadoras, agencias inmobiliarias y portales inmobiliarios (*TINSA, Sociedad de Tasación, Idealista...*) presentan con regularidad índices de precios medios basados en sus propios registros, así como series temporales, estadísticas y gráficas de evolución.

Tal y como expone el investigador José María Montero en varias de sus investigaciones, a un potencial comprador, a un constructor, a un promotor o, en general, a un ciudadano interesado por las cuestiones del mercado de la vivienda, ¿le sirve de algo conocer un precio tan genérico como el que hace referencia a toda una capital de provincia o a toda una Comunidad Autónoma? ¿O una tasa de variación, como es el caso del IPV elaborado por el INE? ¿Es útil para las autoridades con responsabilidad fiscal en materia inmobiliaria? ¿Ayuda en el éxito potencial de la política de vivienda? La respuesta es negativa. Dicho de otro modo, ¿la metodología utilizada para llevar a cabo tales estimaciones y los cálculos de los citados índices es estadísticamente correcta? Quizá la clave reside en que ninguno de los indicadores de precios de vivienda que se elaboran en España tiene en cuenta la Primera Ley de la Geografía (o *Principio de Autocorrelación Espacial*) formulada por Tobler en 1970: "*Todas las cosas están relacionadas entre sí, pero las cosas más próximas están más relacionadas que las que están más distantes*". Así, a medida que las investigaciones sugieren e indican que la autocorrelación espacial se da en la formación del precio de la vivienda, se debiera incluir en la metodología y en el cálculo de índices para la creación de las bases de datos los efectos espaciales que afectan a dicha variable.

## Base de datos utilizada en la investigación

Dado que la investigación del presente trabajo se centra en el estudio del precio de la vivienda en la ciudad de Madrid, la principal fuente de información nos debía proporcionar datos suficientes como para analizar con rigor la formación de los precios de la vivienda en dicha ciudad y, además, que fuesen datos verídicos y actualizados.

De entre todas las fuentes que ofrecen datos de precios de vivienda optamos por la utilización de la base de datos de la página web *Idealista.com*, el portal líder en España para la compraventa de inmuebles.



El portal *Idealista.com* ofrece al usuario una gran variedad de oferta de inmuebles tanto en venta como en régimen de alquiler, y de ello nos hemos servido para la realización de este Trabajo Fin de Máster, ya que uno de los servicios ofrecidos por el portal es ofrece la posibilidad de



[idealista API](#)

Toma de contacto con nuestra base de datos, para poder trastear con nuestros inmuebles

acceder a su base de datos y descargar la información necesaria. Se trata del servicio gratuito “LABS” (algo así como su “laboratorio” de datos), mediante el cual se accede a la base de datos del portal mediante una autorización previa y a través de una contraseña que hay que solicitar al portal. Así, a través de la página web [www.idealista.com](http://www.idealista.com) accedemos al servicio IDEALISTA/LABS ([www.idealista.com/labs](http://www.idealista.com/labs)) e iremos al apartado de API, donde solicitaremos la API KEY que nos remitirán por correo electrónico y que nos dará acceso durante un determinado periodo de tiempo a la base de datos. Además nos remitirán un documento titulado “*Idealista API documentation*” donde explica detalladamente los pasos a seguir y el tipo de información que podemos obtener del portal.

# idealista/labs

La solicitud de la API KEY se realiza enviando un correo a la dirección [labs@idealista.com](mailto:labs@idealista.com) en el que deberemos explicar el propósito de nuestra descarga de información así como informar de la dirección IP del ordenador desde el que accederemos a dicha información.

Tras recibir la API KEY solicitada deberemos proceder a la elaboración de la URL (dirección web) necesaria para acceder a los datos. Para la obtención de datos de este trabajo la API KEY solicitada ha sido “*ehBTBOHk9bjXfhDvVBi1CRdXNRQGzji*” de manera que nuestra URL comienza del siguiente modo: <http://idealista-prod.apigee.net/public/2/search>.

A partir de aquí, siguiendo las indicaciones de la documentación recibida incorporaremos una serie de parámetros a la URL en función de la información que queramos obtener, y una vez completa la dirección web, obtendremos del portal una serie de información en formato JSON en la que aparece toda la información solicitada (se realiza un filtro de las características de las viviendas que necesitamos para la investigación). No existe un orden concreto de colocación de estos parámetros y la información devuelta por el portal de datos nos ofrece datos de 50 viviendas como máximo por página solicitada, con lo que deberemos ir cambiando

el número de página uno a uno hasta conseguir todos los datos que hayamos deseado obtener. Cada parámetro irá precedido del símbolo & por indicaciones de la documentación.

Al final de cada consulta, en la parte baja de la página podremos leer: "*Comprar, De todos los precios, De todos los tamaños*", "*total":200,"totalPages":4,"upperRangePosition":50}* lo que significa que en una consulta con los parámetros establecidos el portal nos devuelve 200 observaciones y, al haber 50 observaciones por página, necesitaremos ir cambiando el parámetro *&numPage* del 1 al 4.

A continuación se exponen los parámetros utilizados en la URL del trabajo (aunque existen otros parámetros opcionales no utilizados) y el proceso seguido para la obtención de los datos, los cuales, una vez obtenidos en el formato JSON, se copiaron y pegaron en un documento *Word (.doc)* para posteriormente hacer lo propio en un archivo *Excel (.xls)* y poder disponer de los datos en un formato viable para el análisis y la realización de los cálculos.

Además de parámetros añadidos, la aplicación devuelve siempre la información relativa a:

- **Size:** tamaño del inmueble
- **Rooms:** número de habitaciones
- **Bathrooms:** número de baños
- **Floor:** planta del inmueble (en caso de encontrarse en un edificio en altura)
- **Address:** dirección del inmueble
- **Price:** precio de la vivienda

Además de otros variados parámetros no utilizados en esta investigación que ayudan a filtrar la información y a obtener los datos del portal de un modo selectivo. A modo de ejemplo, algunos de estos parámetros son características propias del inmueble como *Elevator* (si la vivienda tiene ascensor o no) o *SwimmingPool* (si la vivienda tiene piscina o no) y otros parámetros sirven para filtrar los datos ofrecidos por la web en función de las preferencias, por ejemplo, a la hora de buscar piso de alquiler, como pueden ser *NoSmokers* (compañeros de piso fumadores o no) o *GayPartners* (compañeros de piso gays o no).

**Tabla 1: Parámetros a especificar en la URL para la obtención de datos.**

Parámetro	Descripción
<b>http://idealista-prod.apigee.net/public/2/search</b>	Inicio de la URL (dirección web).
<b>&amp;apikey</b>	La que suministra el portal previa solicitud y registro de la IP.
<b>&amp;operation</b>	Tipo de operación a la que está sujeta la vivienda. Las opciones son A (alquiler) y V (venta). En nuestro caso optamos por las viviendas en venta (V).
<b>&amp;numPage</b>	Representa la página que se está consultando. Debido a que por cada consulta nos devuelve 50 observaciones, será necesario consultar las distintas páginas de una misma consulta.

<b>&amp;center</b>	Las coordenadas (latitud y longitud) del centro del círculo de barrido que marca la superficie desde las que se obtienen los datos. En este trabajo se utilizaron varios centros ya que se establecieron varios círculos de barrido, como se puede observar en las imágenes posteriores.
<b>&amp;distance</b>	La distancia en metros del radio del círculo de barrido que marca la superficie desde las que se obtienen los datos.
<b>&amp;propertyType</b>	Representa el tipo de vivienda solicitado en la búsqueda: piso (flat), estudio (studio), ático (penthouse), chalé (chalet) y dúplex (dúplex).
<b>&amp;country</b>	Indica el país de la búsqueda (España)
<b>&amp;maxItems</b>	Indica el número máximo de observaciones a devolver por cada consulta/página (el máximo permitido es 50)

**30 Tabla de parámetros para la obtención de datos; fuente: elaboración propia a partir de Idealista.**

De este modo, nuestra URL (una de ellas, porque se han utilizado 5 distintas) queda configurada de la siguiente manera:

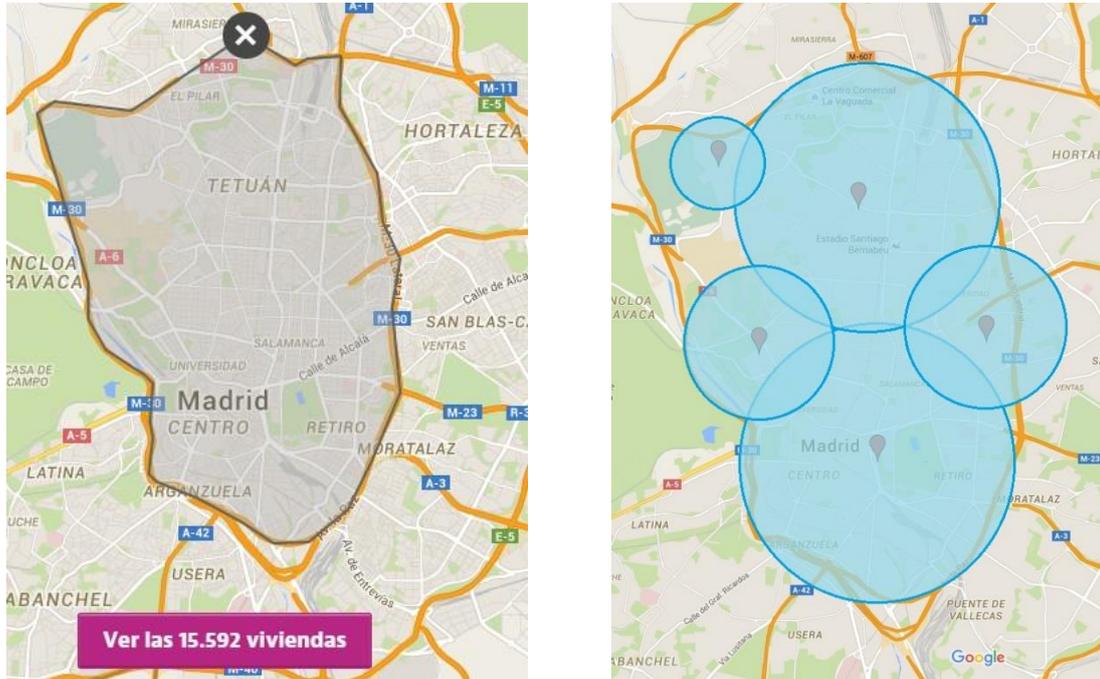
<http://idealista-prod.apigee.net/public/2/search?&apikey=ehBTB0Hk9bjXfhDvVBi1CRdXNRQGzji&country=es&maxItems=50&operation=V&numPage=1&center=40.4293879,-3.7097567&distance=727>

Y a continuación se puede observar una captura de la pantalla de información que devuelve la aplicación una vez introducida dicha URL.



**31 Captura de información para la extracción de la base de datos; fuente: web de Idealista.**

Mediante esta técnica realizamos la operación 5 veces para barrer toda el área correspondiente al Madrid de dentro de la M30, obteniendo más de 15.000 observaciones, como se puede observar en las imágenes adjuntas. Tras eliminar observaciones duplicadas (por el solapamiento de círculos) y quedarnos solamente con las viviendas que ofrecían su dirección exacta (no todas las viviendas que los usuarios suben a la web ofrecen su dirección con exactitud), se obtuvo una base de datos de viviendas consistente en 4.100 observaciones.



32 Área de análisis y extracción de la base de datos del TFM dentro de la M30; fuente: Idealista.com



33 Distritos del municipio de Madrid; fuente: munimadrid.org

## Fuentes de datos sobre elementos georreferenciados

Dado que la investigación del presente Trabajo Fin de Máster trata de establecer la influencia de factores espaciales de cercanía y lejanía a ciertos elementos en la formación de los precios de la vivienda resultaba necesario obtener información georreferenciada de dichos elementos, es decir, coordenadas geográficas. Para ello recurrimos a la web de datos abiertos del Ayuntamiento de Madrid y a la propia georreferenciación manual de elementos en la aplicación *Maps* de *Google* y mediante el programa *ARCGIS*, un software SIG (Sistemas de Información Geográfica, GIS por sus siglas en inglés). Información más detallada sobre la georreferenciación se puede obtener en el apartado anterior del trabajo dedicado a los Sistemas de Información Geográfica).

En cuanto al portal de datos del Ayuntamiento, según se explica en la propia página web <http://datos.madrid.es/portal/site/egob>, *Datos Abiertos (open data en inglés) es una iniciativa global, ligada a las políticas de Gobierno Abierto, que persigue que los datos y la información, especialmente las que poseen las administraciones públicas, se publiquen de forma abierta, regular y reutilizable para todo el mundo, sin restricciones de acceso, copyright, patentes u otros mecanismos de control.*

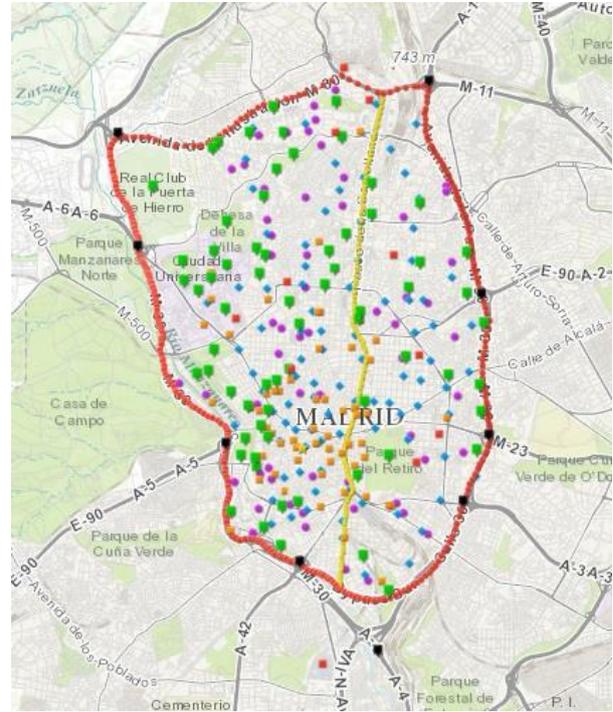
Del portal de Datos Abiertos y de la georreferenciación manual de algunos elementos se ha elaborado una base de datos georreferenciada con las coordenadas geográficas (latitud y longitud) de los siguientes elementos:

- Colegios
- Hospitales
- Parques y Jardines
- Estaciones de Metro
- Museos
- Kilómetro cero: Puerta del Sol
- M30: vía de circunvalación de la ciudad de Madrid
- Paseo de la castellana: eje Norte-Sur de la ciudad
- Nudos de comunicación en la M30

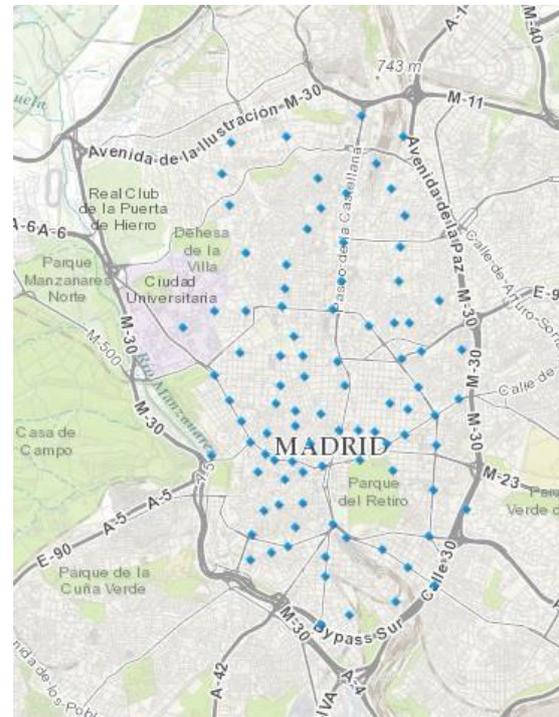


34 Web de Datos Abiertos del Ayuntamiento de Madrid; fuente: web de [datos.madrid.es](http://datos.madrid.es).

A continuación se muestran imágenes de algunos de los mapas utilizados para georreferenciar elementos, tanto en la aplicación *Maps* de *Google* como en *ARCGIS*.



**35 Georreferenciación de elementos en Maps y en ARCGIS online; fuente: elaboración propia.**



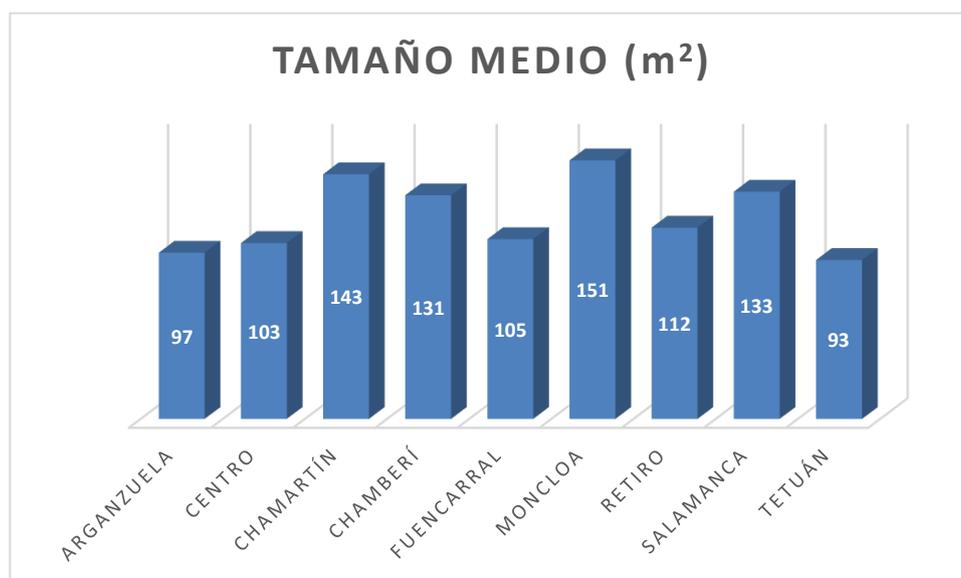
**36 Georreferenciación de colegios y estaciones de metro; fuente: elaboración propia.**

## RESULTADOS

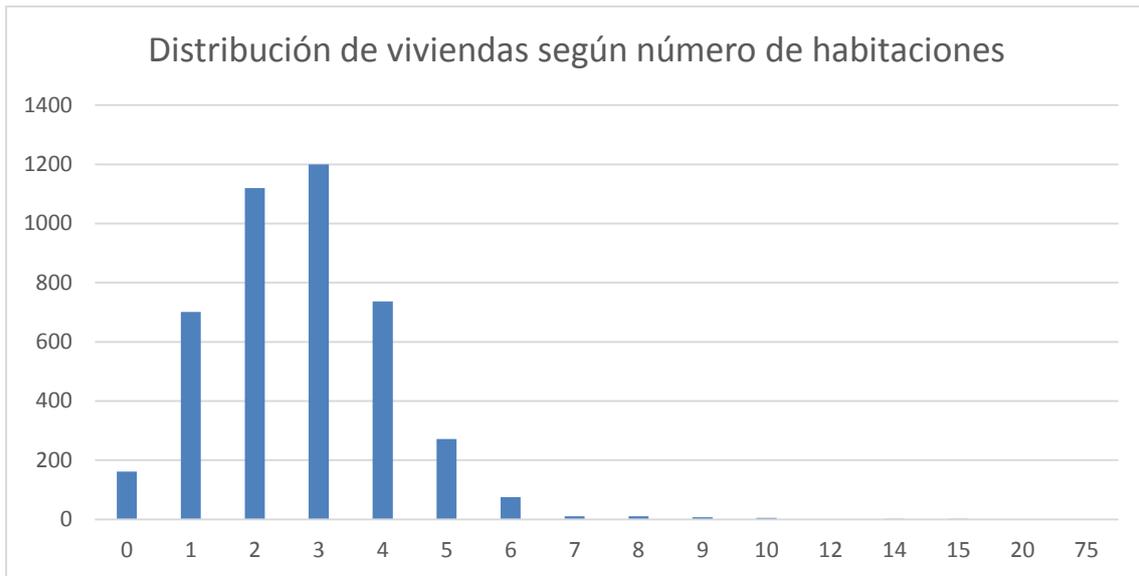
### Análisis descriptivo de los datos

Antes de proceder al análisis de los datos de las 4.100 observaciones que componen la base de datos de este TFM hemos analizado las viviendas en base al barrio y al distrito al que pertenecen. Cabe recordar que en el presente estudio se han analizado las viviendas que se encuentran dentro de la M30, ya que se ha escogido esta vía de circunvalación como límite para el estudio espacial del TFM. Del mismo modo, del análisis descriptivo de las observaciones de la base de datos obtendremos datos estadísticos que nos proporcionan información acerca de las características de las viviendas así como de los precios de las mismas. Al final del trabajo, en el *Anexo 1* se muestran los gráficos pormenorizados con los precios de mercado y los valores catastrales calculados por distritos y por barrios.

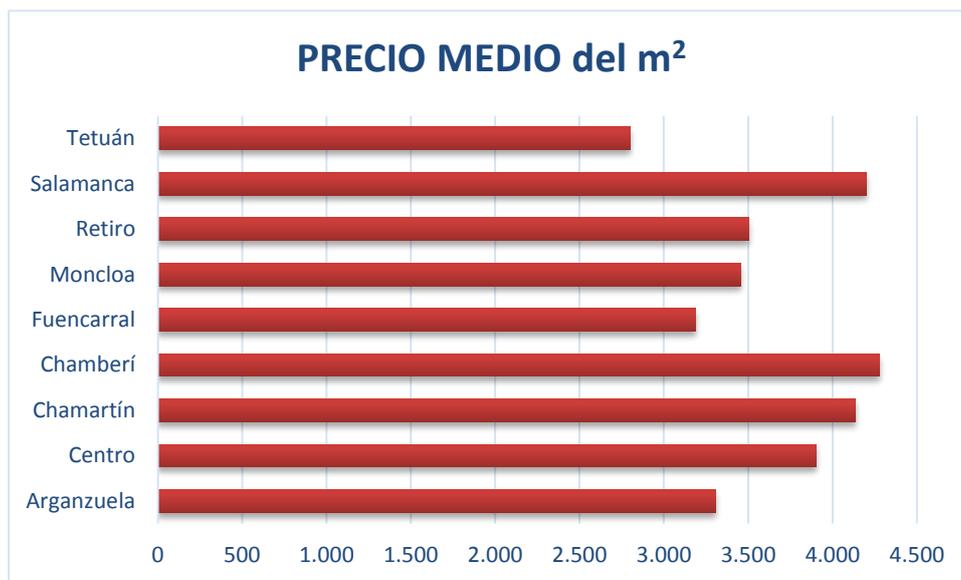
Así, tras el análisis de los datos, tenemos que de las 4.100 viviendas analizadas, el tamaño medio es de 117 m<sup>2</sup>, el precio medio de la muestra es de 433.750 € y el precio medio por metro cuadrado es de 3639 €/m<sup>2</sup>. El promedio del número de habitaciones es de 2,74 y tal como podemos observar en los gráficos que se muestran a continuación, la mayoría de viviendas tienen 2 o 3 habitaciones, aunque también hay muchas con 1 y con 4. En los siguientes gráficos se muestra información del análisis descriptivo de la muestra analizada según los distritos estudiados.



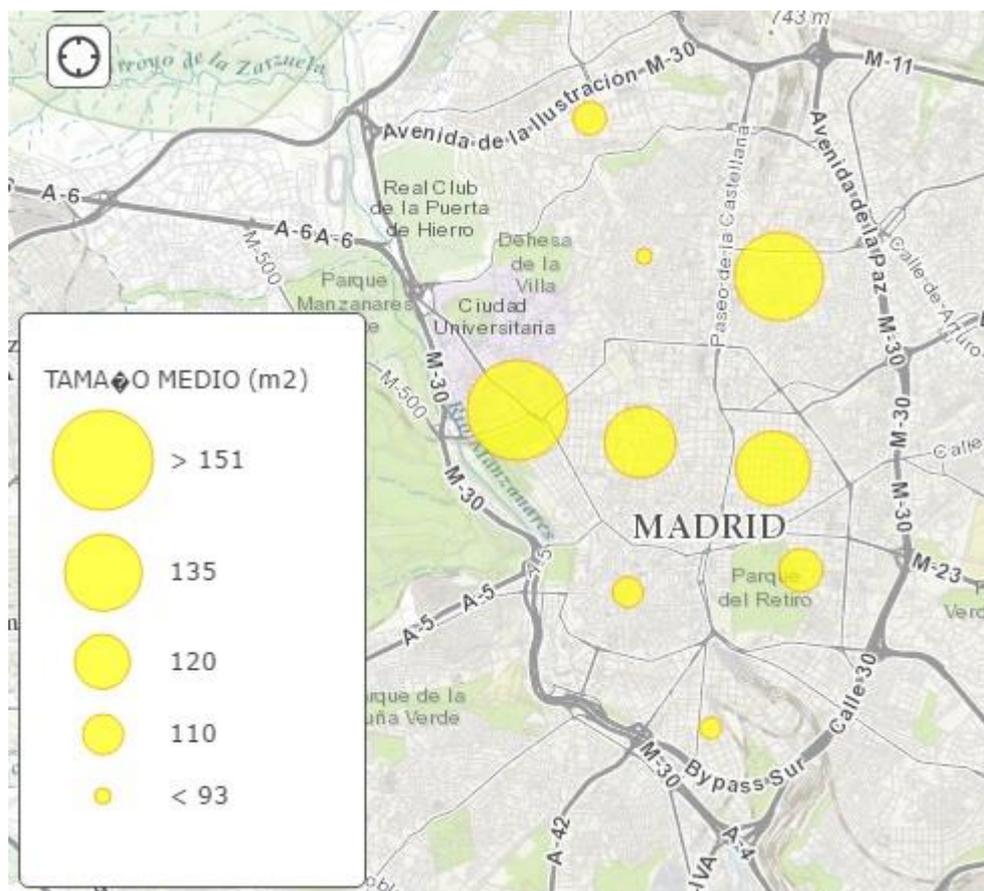
37 *Tamaño medio de las viviendas por distritos; fuente: elaboración propia.*



**38** *Número de habitaciones de las viviendas por distritos; fuente: elaboración propia.*



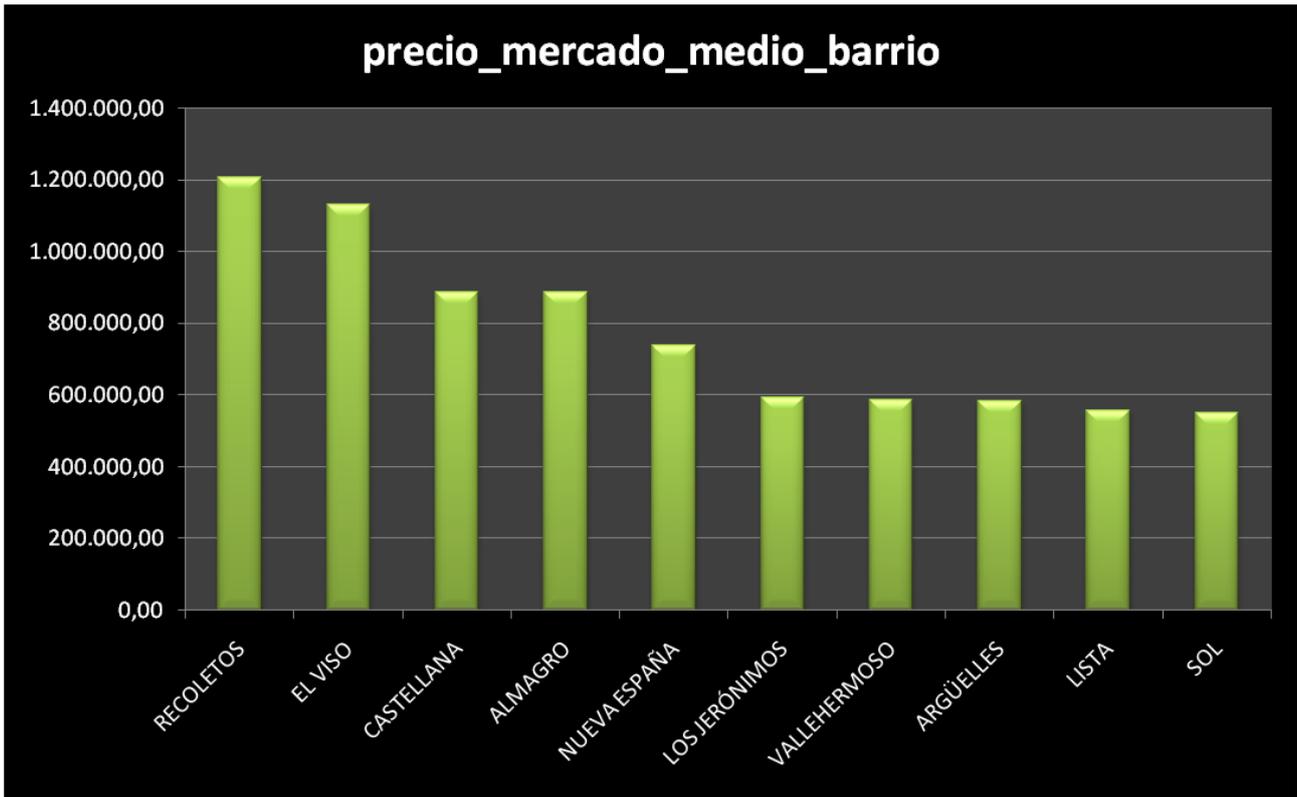
**39** *Precio medio de las viviendas por distritos; fuente: elaboración propia.*



40 Tamaño medio de las viviendas analizadas por distritos; fuente: elaboración propia.

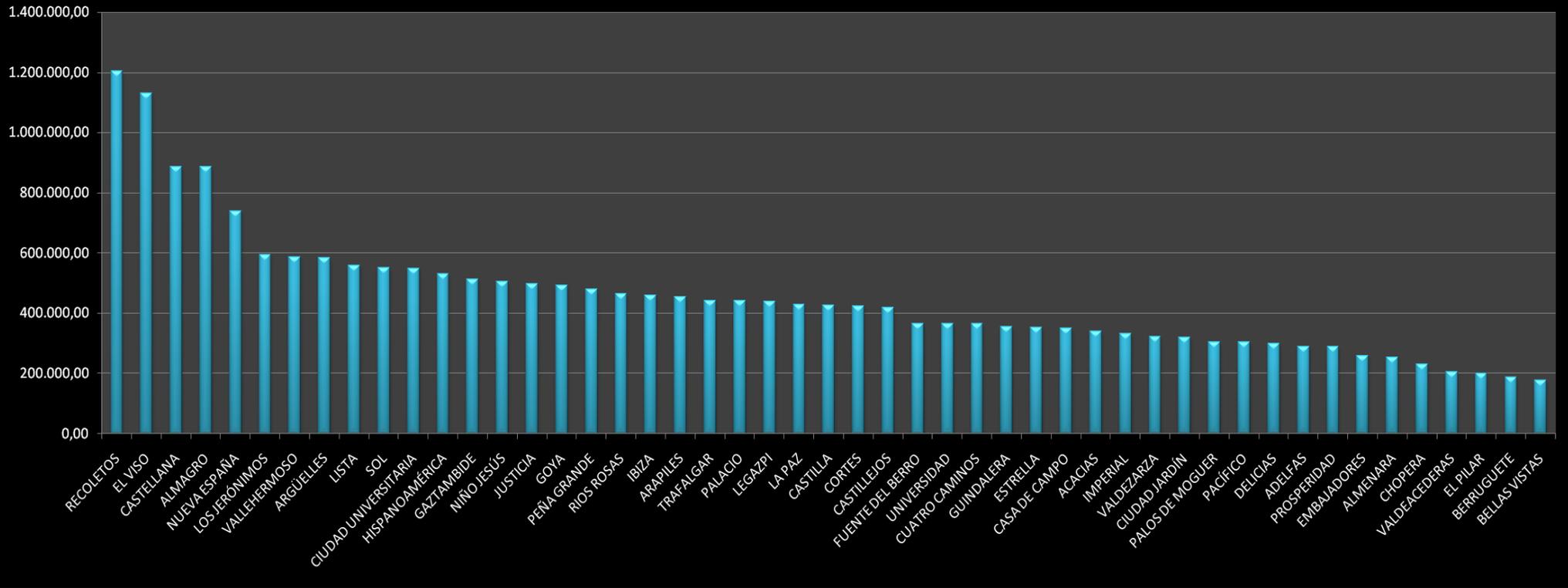
DISTRITO	TAMAÑO MEDIO (m <sup>2</sup> )	PRECIO MEDIO MERCADO	VALOR MEDIO CATASTRAL	PRECIO MEDIO MERCADO del m <sup>2</sup>
Arganzuela	97	312.530	123.696	3.304
Centro	103	398.169	163.807	3.901
Chamartín	143	627.347	221.884	4.136
Chamberí	131	542.830	184.790	4.278
Fuencarral	105	344.191	161.289	3.187
Moncloa	151	513.460	259.054	3.454
Retiro	112	387.360	217.797	3.502
Salamanca	133	576.618	242.654	4.203
Tetuán	93	259.783	111.283	2.798

41 Tabla de tamaños y precios por distritos; fuente: elaboración propia.



*42 Barrios con precio medio de mercado más alto; fuente: elaboración propia.*

### precio\_mercado\_medio\_barrio



43 Precio medio de mercado de todos los barrios analizados en el trabajo; fuente: elaboración propia.

## Análisis exploratorio de datos espaciales (AEDE)

El índice de Moran calculado es de 0,41.

Al resultar el Índice de Moran positivo, los valores de la muestra de datos (las 4.100 observaciones) tienden a agruparse espacialmente (los valores altos se agrupan cerca de otros valores altos; los valores bajos se agrupan cerca de otros valores bajos), lo cual confirma que existe dependencia espacial significativa al presentar un p-valor de 0,0010 (inferior a 0,05) a 999 permutaciones.

En cuanto al gráfico de cuadrantes, las cuatro áreas del gráfico corresponden a los cuatro tipos de asociación espacial local posibles (entre una vivienda y sus vecinas): viviendas con alto precio (superior a la media) rodeadas de viviendas de alto precio a su vez (cuadrante superior derecho), viviendas de precio bajo rodeadas de viviendas de precio alto (cuadrante superior izquierdo), viviendas de precio bajo rodeadas de viviendas de precio bajo (cuadrante inferior izquierdo) y, finalmente, viviendas con precios altos rodeadas de viviendas de precios bajos (cuadrante inferior derecho).

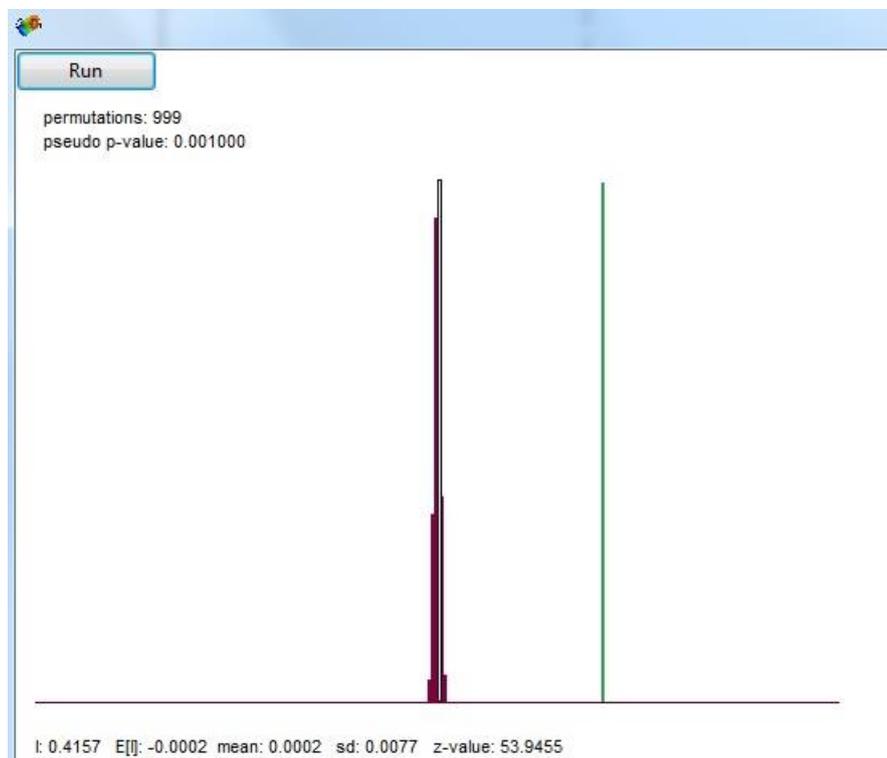
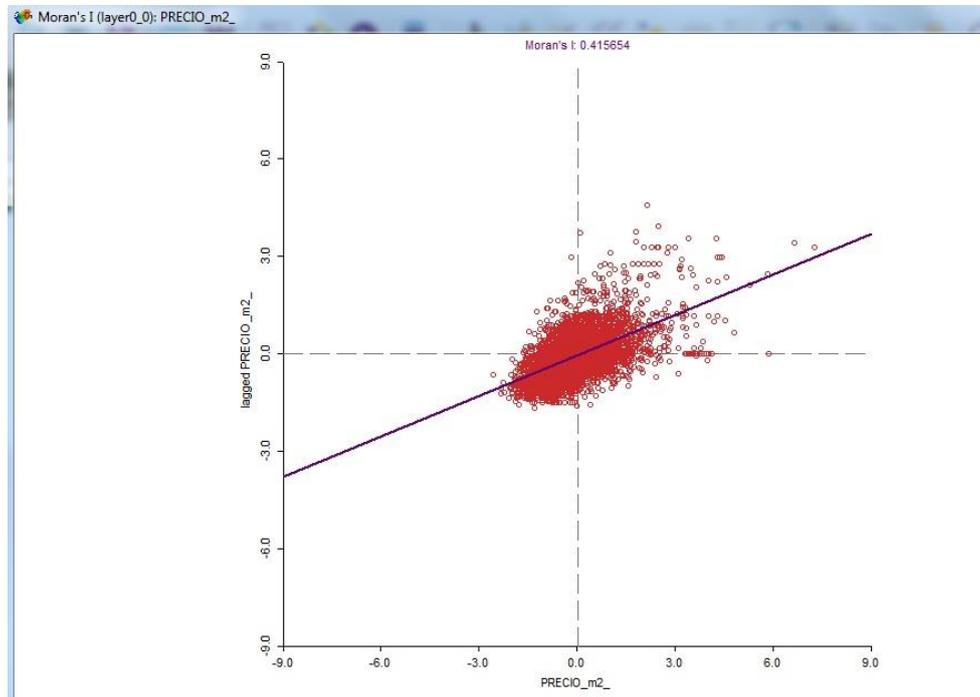
Por tanto, **los cuadrantes I y III corresponden a las formas positivas de autocorrelación espacial, mientras que los II y IV representan la autocorrelación espacial negativa**. Estos dos últimos casos se refieren a las localizaciones que podríamos considerar como atípicas. Las nubes de puntos indican las observaciones que se encuentran en cada caso de los cuatro posibles.

La autocorrelación espacial global (IM) también puede verse en el gráfico como la pendiente de la recta de regresión, en este caso positiva ya que el valor del índice es positivo (0,41).

En cuanto a la autocorrelación local (LISA) observamos en el primer gráfico (*significance map*) la existencia de clústeres o grupos/conjuntos de viviendas que, en cuanto a su precio (variable estudiada en los cálculos), tienden a agruparse. Así, observamos esta correlación espacial en varias zonas: al norte en el distrito de Tetuán, al sur en el de Arganzuela, y en la zona centro (aunque de menor intensidad) en los barrios de Almagro y en los distritos de Chamartín y Salamanca, especialmente en la zona del Paseo de la Castellana.

Para entender el significado de estas agrupaciones acudimos al segundo mapa (*LISA cluster map*) en el que observamos **en color azul marino la existencia de clústeres low-low** (viviendas de precio bajo rodeadas de viviendas de precio bajo) en los distritos de Arganzuela (sur) y en Tetuán (norte), mientras que **de color rojo se distinguen las asociaciones de viviendas high-high** (viviendas de precio alto rodeadas de viviendas de precio alto) de los distritos de Chamberí, Salamanca y Chamartín, en los que se observa la tendencia de éstas últimas a agruparse en torno al eje norte-sur que supone el Paseo de la Castellana.

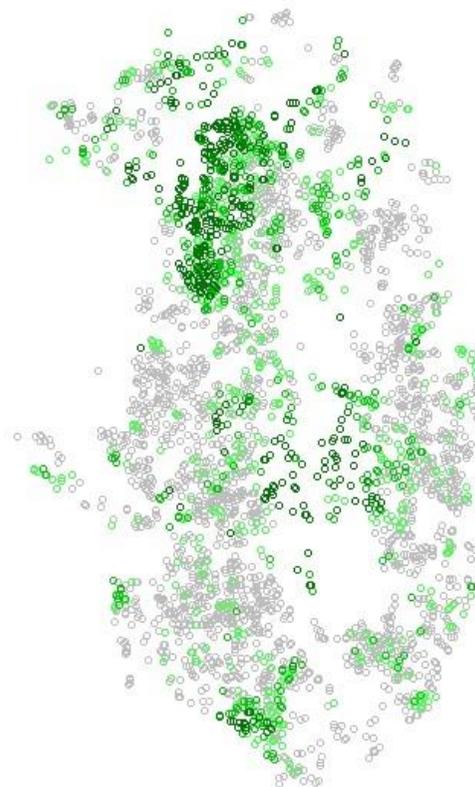
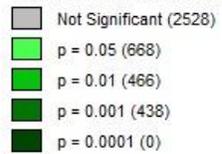
## Autocorrelación espacial global (índice de Moran)



44 Índice de Moran; fuente: elaboración propia con el software GeoDa.

## Autocorrelación espacial local (Local Indicator of Spatial Association -LISA-)

LISA Significance Map: layer0\_0, I\_PRECIO\_m2\_ (999 perm)



45 LISA significance map; fuente: elaboración propia con el software GeoDa.

LISA Cluster Map: layer0\_0, L\_PRECIO\_m2\_ (999 perm)

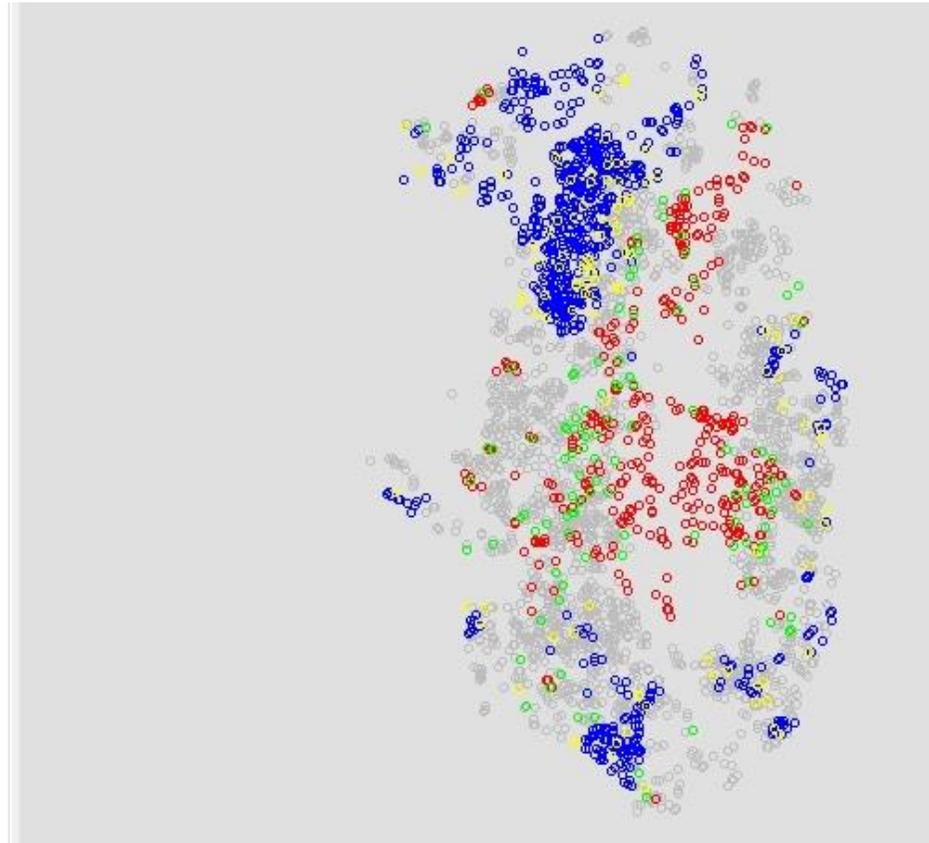
Not Significant (2528)

High-High (466)

Low-Low (858)

Low-High (139)

High-Low (109)



46 LISA cluster map; fuente: elaboración propia con el software GeoDa.

Resultados de las regresiones y análisis de coeficientes

TABLA 2: Resultados de las regresiones

Variable	Modelo 1			Modelo 2		
	Coef.	Error típ.	t-valor	Coef.	Error típ.	t-valor
[constante]	5.079,51	18.768,5	0,270641	-53.104,968*	8.438,928	-6,293
[planta]	10.373,7*	1.106,36	9,37636	10.884,351*	1.146,780	9,491
[exterior]	19.780,6*	6.655,01	2,97229	12.250,467***	6.792,461	1,804
[tamaño (m <sup>2</sup> )]	3.367,04*	61,9874	54,3181	3.619,963*	60,900	59,441
[habitaciones]	-41.045,8*	2.224,24	-18,4538	-45.653,197*	2.262,806	-20,175
[baños]	65.671*	4.311,79	15,2306	69.483,657*	4.408,991	15,760
[chalé]	415.046*	156.044	2,65981	318.685,759**	159.627,680	1,996
[dúplex]	-32.731,5**	13.605,2	-2,40581	-48.385,249*	13.971,236	-3,463
[estudio]	-103.853*	13.941,1	-7,44943	-105.811,849*	143.13,297	-7,393
[ático]	32.728,5*	10.935,1	2,99298	31.668,734*	11.224,125	2,821
[distancia al km 0 (Puerta del Sol)]	-34.976,6*	3.977,46	-8,7937	-	-	-
[distancia al km 0 (Puerta del Sol) < 2 km]	-	-	-	61.790,118*	6.904,638	8,949
[2 km < distancia al km 0 (Puerta del Sol) < 3 km]	-	-	-	63.521,812*	7.052,911	9,006
[distancia al Paseo de la Castellana (eje norte-sur)]	-22.118*	4.149,28	-5,33056	-	-	-
[distancia al Paseo de la Castellana (eje norte-sur) < 0,3 km]	-	-	-	12.284,028	9.389,595	1,308
[0,3 km < distancia al Paseo de la Castellana (eje norte-sur) < 0,7 km]	-	-	-	38.391,984*	7.951,000	4,829
[distancia a la M30]	26.207*	8.321,39	3,14935	-	-	-
[distancia a la M30 < 0,1 km]	-	-	-	7.767,407	21.666,561	0,358
[0,1 km < distancia a la M30 < 0,3 km]	-	-	-	-29.714,519*	11.265,432	-2,638
[distancia a los nudos de la M30]	-16.817,4*	6.422,86	-2,61837	-	-	-
[distancia a los nudos de la M30 < 0,4 km]	-	-	-	-41.159,51	30.283,481	-1,359
[0,4 km < distancia a los nudos de la M30 < 0,8 km]	-	-	-	-59.279,856*	11.137,796	-5,322
[distancia a las estaciones de metro]	-4.207,14	13.739,7	-0,306204	-	-	-
[distancia a las estaciones de metro < 0,05 km]	-	-	-	52.704,179*	18.384,409	2,867
[0,05 km < distancia a las estaciones de metro < 0,1 km]	-	-	-	1.083,443	11.937,658	0,091

**TABLA 2 (continuación): Resultados de las regresiones**

Variable	Modelo 1			Modelo 2		
	Coefficiente	Error típ.	t-valor	Coefficiente	Error típ.	t-valor
[distancia a zonas verdes (parques y jardines)]	-59.192*	9.229,29	-6,4135	-	-	-
[distancia a zonas verdes (parques y jardines) < 0,1 km]	-	-	-	57.818,837**	23.764,004	2,433
[0,1 km < distancia a zonas verdes (parques y jardines) < 0,3 km]	-	-	-	18.039,587*	6.821,643	2,644
[distancia a hospitales]	-18.843*	5.087,23	-3,70399	-	-	-
[distancia a hospitales < 0,3 km]	-	-	-	40.909,833*	12.664,852	3,230
[0,3 km < distancia a hospitales < 0,6 km]	-	-	-	26.302,009*	7.711,551	3,411
[distancia a colegios]	98.128,9*	13.097,4	7,49222	-	-	-
[distancia a colegios < 0,1 km]	-	-	-	-33.325,903*	12.223,950	-2,726
[0,1 km < distancia a colegios < 0,3 km]	-	-	-	-42.766,484*	5.662,516	-7,553
[distancia a museos]	66.665,9*	7.988,72	8,34501	-	-	-
[distancia a museos < 0,1 km]	-	-	-	31.248,261***	17.052,899	1,832
[0,1 km < distancia a museos < 0,3 km]	-	-	-	-11.753,15	7.313,269	-1,607
[weight_4]	9,66492*	3,50634	2,75641	-	-	-
[w_size]	615,49*	56,8888	10,8192	-	-	-
[R <sup>2</sup> corregida del modelo]	0.776	-	-	0.756	-	-

\* Significativo al 1 % (p-valor < 0,01)

\*\* Significativo al 5 % (p-valor < 0,05)

\*\*\* Significativo al 10 % (p-valor < 0,1)

47 *Tabla de resultados de los modelos de regresión; fuente: elaboración propia.*

En los dos modelos de regresión planteados en el trabajo (*modelo 1* y *modelo 2*) se ha trabajado con el método de mínimos cuadrados y se ha establecido la variable “precio del inmueble” como variable dependiente y todas las demás como variables explicativas independientes. Algunas de estas variables independientes responden a un tipo de modelo denominado de “regresión hedónica” en el que se incluyen variables propias del inmueble como su tamaño, el número de habitaciones o baños, la planta en la que se encuentra o si la vivienda es exterior o interior.

En el *modelo 1* se ha planteado una regresión de tipo espacial incorporando dos variables independientes creadas con la finalidad de comprobar la existencia o no de autocorrelación espacial, es decir, si el precio y las características de las viviendas del entorno influyen o no en el precio de cada inmueble. Estas dos variables son *weight\_4* y *w\_size*. La variable *weight\_4* se ha creado con una matriz de pesos que tiene en cuenta las 4 viviendas más cercanas a cada observación, para determinar si hay influencia espacial de estas 4 viviendas sobre cada vivienda en cuestión. De un modo similar, la variable *w\_size* actúa de un modo similar pero en este caso para determinar la influencia del tamaño de las viviendas “vecinas” en el tamaño de cada observación.

Por otro lado, las variables denominadas como “distancia a” pretenden dotar al modelo de una capacidad explicativa que determine la influencia de factores espaciales en la formación del precio de las viviendas, es decir, de la cercanía de las viviendas a determinados lugares o puntos de interés (zonas verdes, colegios, estaciones de metro, etc.). Mientras en el *modelo 1* se han tomado las variables como la distancia euclídea total desde cada observación hasta cada elemento estudiado (se ha calculado la distancia de cada observación hasta el elemento en cuestión más cercano), en el *modelo 2* se han segmentado cada una de estas distancias para establecer “radios de influencia” y determinar qué radio influye y de qué manera en la formación del precio de las viviendas. De este modo, en el *modelo 2* se han tomado como variables independientes las distintas distancias a los elementos estudiados en el trabajo (colegios, zonas verdes, hospitales, estaciones de metro...), pero esta vez habiendo realizado una segmentación de las distancias para comprobar el efecto de cada variable sobre el precio en un determinado radio. Para ello se han convertido todas las distancias euclídeas calculadas en variables dicotómicas y estudiando el efecto sobre la variable dependiente de cada variable en varios segmentos (áreas circulares con un determinado radio).

En ambos modelos se han obtenido valores de la  $R^2$  corregida de 0,776 (*modelo 1*) y de 0,756 (*modelo 2*), lo cual indica que el *modelo 1* explica aproximadamente el 78% de la formación del precio de la vivienda (variable dependiente) mientras que el *modelo 2* hace lo propio en un 76%.

A continuación se exponen los resultados de las regresiones realizadas en los dos modelos así como el análisis de los coeficientes obtenidos de las mismas. La unidad en la que se miden los coeficientes es en €.

- La **constante** presenta un valor de 5.079,51 en el *modelo 1* y de -53.104,968 en el *modelo 2*, siendo solamente significativa en el *modelo 2*. En cualquier caso, este valor sería el precio de una vivienda en el caso de que no presentara ninguna de las características que aportan el resto de variables analizadas.

- La variable **planta**, que adopta los valores de 10.373,7 y de 10.884,351, es significativa en ambos modelos y representa que por cada planta de altura en que se encuentra el inmueble el precio se incrementa en dichas cantidades.

- La variable **exterior** adopta los valores 19.780,6 y 12.250,467 pero sólo es significativa en el *modelo 1*, lo que significaría que el hecho de que la vivienda fuese exterior incrementaría el precio de la vivienda en 19.780,6 €.

- La variable **tamaño** adopta los valores de 3.367,04 en el *modelo 1* y de 3.619,963 en el *modelo 2*, siendo además significativos en ambos modelos. Estos valores indican que por cada metro cuadrado que aumenta la vivienda el precio aumenta en dichas cantidades, lo cual es coherente con el precio medio del metro cuadrado en la ciudad de Madrid en la actualidad, que ronda los 3.500 €.

- La variable **habitaciones**, que adopta los valores de -41.045,8 y de -45.653,197, es significativa en ambos modelos y el hecho de que este coeficiente sea negativo en ambos modelos se explica en que el factor tamaño ya se ha introducido en la regresión en la variable “tamaño”, y el hecho de que una vivienda tenga más habitaciones hace que éstas sean más pequeñas y ello redunde en una disminución del precio del inmueble.

- La variable **baños** adopta los valores de 65.671 en el *modelo 1* y de 69.483,657 en el *modelo 2*, siendo además significativos en ambos modelos. Estos valores indican que por cada baño adicional el precio de la vivienda se incrementa en dichas cantidades.

- La variable **chalé** adopta los valores de 415.046 en el *modelo 1* y de 318.685,759 en el *modelo 2*, siendo los coeficientes significativos en ambos modelos. Esto indicaría que el hecho de que una vivienda sea un chalé incrementa el precio en dichas cantidades, pero en el caso de esta variable hay que tener en cuenta que de las 4100 observaciones de la base de datos de las regresiones sólo hay 1 chalé, con lo cual no es un coeficiente que explique de manera significativa la formación de la variable precio, al tratarse de una única observación entre las 4100 de la base de datos analizada.

- La variable **dúplex**, que adopta los valores de -32.731,5 en el *modelo 1* y de -48.385,249 en el *modelo 2*, presenta en ambos casos coeficientes negativos ya que los dúplex suelen ser más grandes que los pisos, y por ello puede ser que la variable independiente “tamaño” sea de nuevo la que explique la variable dependiente “precio”.

- La variable **estudio**, que adopta los valores de -103.853 en el *modelo 1* y de -105.811,849 en el *modelo 2*, presenta también en ambos casos coeficientes negativos. Los estudios son inmuebles generalmente más pequeños que los pisos y puede que en este caso, de nuevo, la variable “tamaño” explique la reducción del precio.

- La variable **ático** adopta los valores de 32.728,5 en el *modelo 1* y de 31.668,734 en el *modelo 2*, siendo los coeficientes significativos en ambos modelos. Esto indicaría que el hecho de que una vivienda sea un ático incrementaría el precio en dichas cantidades. En este caso es coherente que ambos coeficientes sean de signo positivo en ambos modelos ya que, por norma general, los áticos suelen ser inmuebles más caros que los pisos de plantas inferiores por las características propias de un ático (mejores vistas, más luz, posición “privilegiada” en el

edificio, etc.), pero en este caso hay que tener en cuenta que las variables “tamaño” y “planta” también explicarían de manera significativa la variable “precio”.

A continuación se exponen los resultados y el análisis de los coeficientes de las variables correspondientes a las distancias, teniendo en cuenta que éstas están medidas en kilómetros, con lo que, en la regresión del *modelo 1*, el aumento o disminución del precio será por kilómetro de distancia y un coeficiente positivo implicará un aumento de precio cuanto más lejos se encuentre el inmueble del elemento estudiado mientras que un coeficiente negativo implicará aumento de precio cuanto más cerca se encuentre el inmueble del elemento estudiado. En la interpretación de estos resultados hay que tener en cuenta que la variable precio no aumenta ni disminuye de manera lineal a medida que aumenta o disminuye la distancia, por lo que estos resultados nos sirven para obtener una idea general de cómo se comportan espacialmente los precios en la zona de la ciudad de Madrid analizada, no pudiendo realizarse relaciones lineales entre la variable precio y las variables con efectos espaciales (como las distancias segmentadas).

- La variable **distancia al km 0 (Puerta del Sol)** presenta un coeficiente de valor -34.976,6 en el *modelo 1*, significativo al 1%, con lo que refleja un aumento de precio de 34.976,6 € por cada km de cercanía a la Puerta del Sol.

En el *modelo 2*, en el que se han segmentado las distancias, el coeficiente es de 61.790,118 para una distancia inferior a 2 km y de 63.521,812 para la zona comprendida entre los 2 km y los 3 km. Los coeficientes del *modelo 2* indicarían que, teóricamente, los precios de la vivienda en esas zonas se incrementarían la cantidad correspondiente a sus coeficientes.

- La variable **distancia al Paseo de la Castellana (eje norte-sur)** presenta un coeficiente de valor -22.118 en el *modelo 1*, significativo al 1%, con lo que refleja un aumento de precio de 22.118 € por cada km de cercanía a esta significativa avenida que representa el principal eje norte-sur de la ciudad.

En el *modelo 2*, el coeficiente es de 38.391,984 para la zona comprendida entre los 300 m y los 700 m, lo que indica un aumento de precio total por dicho valor en esa determinada franja. Este resultado es coherente con los resultados obtenidos en los mapas de georreferenciación realizados con ARCGIS, donde se observa con claridad que las viviendas cercanas a la Castellana son las que presentan los precios más elevados de la ciudad en términos de precio del metro cuadrado. Además, los precios más altos se dan en barrios como Almagro, El Viso o Castellana, barrios que se articulan a lo largo del eje que configura esta vía.

- La variable **distancia a la M30** presenta un coeficiente de valor 26.207 en el *modelo 1*, significativo al 1%, con lo que refleja un aumento de precio de dicha cantidad por cada km de lejanía a dicha vía, o lo que es lo mismo, una disminución de precio de 26.207 € por cada km de cercanía a esta significativa e importante circunvalación (que en realidad se llama Calle30).

En el *modelo 2*, el coeficiente es de -29.714,519 para la zona comprendida entre los 100 m y los 300 m, lo que indica una disminución de precio total por dicho valor en esa determinada franja.

- La variable **distancia a los nudos de la M30** presenta un coeficiente de valor -16.817,4 en el *modelo 1*, significativo al 1%, con lo que refleja un aumento de precio de dicha cantidad por cada km de cercanía a estos puntos. Cabe recordar que los nudos elegidos para el estudio de esta variable son los correspondientes a las principales salidas de las grandes vías de la ciudad, los cuales forman importantes nudos de comunicación en sus respectivos cruces con la M30. En una gran ciudad como Madrid, donde los transportes y la comunicación en coche y en transporte público es de suma importancia, el significado de este coeficiente se puede explicar atendiendo a que se valora el hecho de vivir cerca de estos nudos (que son las salidas de la A1, A2, A3..., y que son la vía de salida y de entrada para muchos madrileños en su camino diario hacia sus respectivos puestos de trabajo).

En el *modelo 2*, el coeficiente es de -59.279,856 para la zona comprendida entre los 400 m y los 800 m, lo que indica una disminución de precio total por dicho valor en esa determinada franja. En cambio, el coeficiente para la zona de influencia inferior a los 400 m no es significativo.

- La variable **distancia a las estaciones de metro** presenta un coeficiente de valor -4.207,14 en el *modelo 1*, pero hay que destacar que es la única variable espacial que no presenta un coeficiente significativo en este modelo. El significado de este valor nos indicaría, en caso de haber sido significativo, un aumento de precio de 4.207,14 € por cada km de cercanía a una estación de metro.

En el *modelo 2*, el coeficiente es de 52.704,179 para las viviendas situadas en una zona inferior a 50 m de una parada de metro. Este coeficiente sí que es significativo al 1%, e indica un aumento del precio total por dicho valor en esa determinada área circular de 50 m de radio alrededor de cada estación de metro.

- La variable **distancia a zonas verdes (parques y jardines)** presenta un coeficiente, significativo al 1%, de valor -59.192 en el *modelo 1*, con lo que refleja un aumento de precio de dicha cantidad por cada km de cercanía a una zona verde, ya sea parque o jardín de cierta entidad.

En el *modelo 2*, el coeficiente es de 57.818,837, significativo al 5%, para la zona de viviendas situadas a menos de 100 m de una zona verde, mientras que para la zona comprendida entre los 100 m y los 300 m el coeficiente es de 18.039,587 y significativo al 1%, lo que indica un aumento del precio total de las viviendas por valor de dichas cantidades en esas determinadas zonas.

- La variable **distancia a hospitales** presenta un coeficiente de valor -18.843 en el *modelo 1*, significativo al 1%, con lo que refleja un aumento de precio de dicha cantidad por cada km de cercanía a un hospital.

En el *modelo 2*, el coeficiente es de 40.909,833 y significativo al 1% para la zona de viviendas situadas a menos de 300 m de un hospital, mientras que para la zona comprendida entre los 300 m y los 600 m el coeficiente es de 26.302,009, también significativo al 1%, lo que indica un aumento del precio total de las viviendas por valor de dichas cantidades en esas determinadas zonas.

- La variable **distancia a colegios** presenta un coeficiente de valor 98.128,9 en el *modelo 1*, significativo al 1%, con lo que refleja una disminución de precio de dicha cantidad por cada km de cercanía a un colegio.

En el *modelo 2*, el coeficiente es de -33.325,903 y significativo al 1% para la zona de viviendas situadas a menos de 100 m de un colegio, mientras que para la zona comprendida entre los 100 m y los 300 m el coeficiente es de -42.766,484, también significativo al 1%, lo que indica una bajada del precio total de las viviendas por valor de dichas cantidades en esas determinadas zonas.

- La variable **distancia a museos** presenta un coeficiente de valor 66.665,9 en el *modelo 1*, significativo al 1%, con lo que refleja una disminución de precio de dicha cantidad por cada km de cercanía a un museo.

En el *modelo 2*, ninguno de los coeficientes obtenidos es significativo.

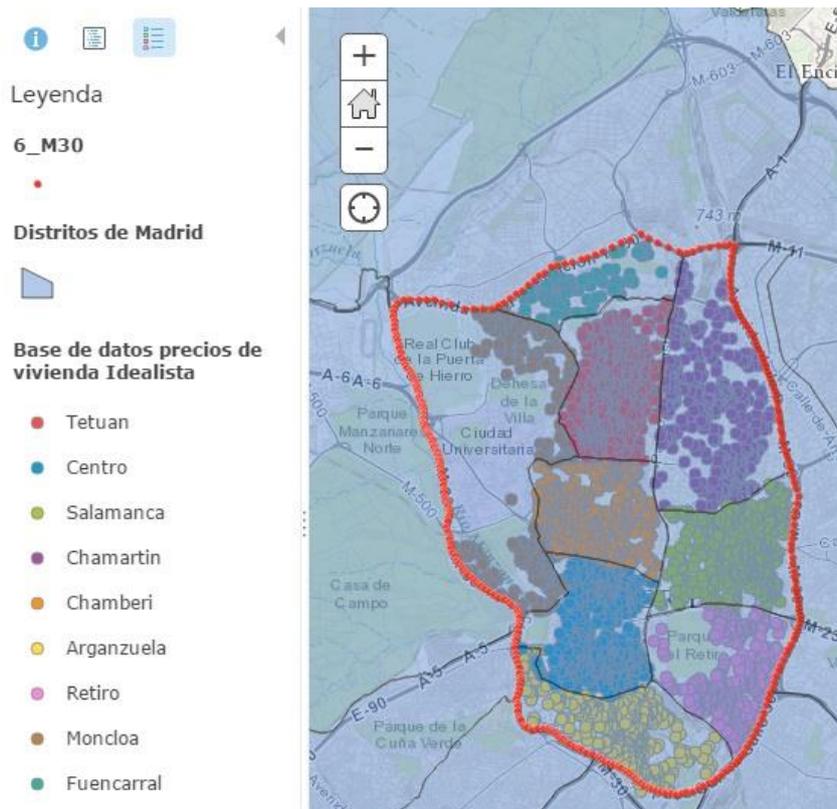
Los resultados del *modelo 1* correspondientes a la regresión espacial son los siguientes:

- La variable **weight\_4** presenta en el *modelo 1* un valor de 9,66492 significativo al 1%.

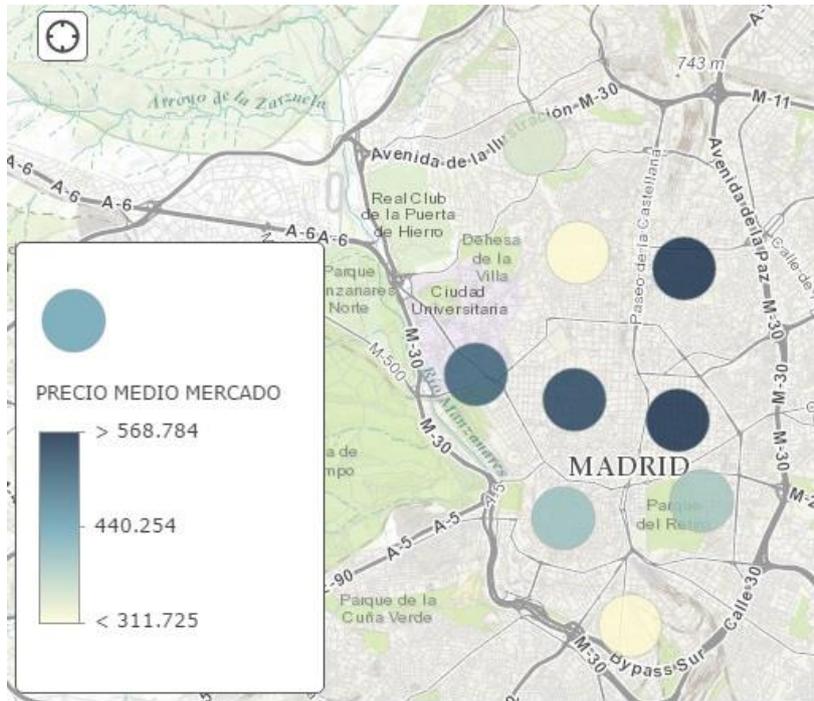
- La variable **w\_size** presenta en el *modelo 1* un valor de 615,49 significativo al 1%.

## Resultados y análisis de los mapas de *ARCGIS*

En este apartado de los resultados se muestran gráficamente los análisis realizados con el software *SIG ArcGis online*, un sistema de información geográfica con el que se ha tratado y analizado la información relativa a las 4.100 observaciones del TFM en lo relativo a sus características *indoor* y *outdoor* (precio, distancias, etc.).



Los sistemas de información geográfica (SIG), como el utilizado en el presente trabajo, permiten georreferenciar elementos para visualizar sus propiedades, características, y las posibles relaciones entre ellos. En este caso, se han referenciado mediante coordenadas de latitud y longitud las 4.100 viviendas correspondientes a la zona de estudio del trabajo (la zona interior a la M30). Además de las propias características de las viviendas analizadas se han georreferenciado los elementos a los que se quería analizar en lo relativo a las distancias de éstos desde cada una de las viviendas y su influencia en la formación del precio de las mismas.



49 Precio medio de mercado de la base de datos analizada por distritos; fuente: elaboración propia.

Uno de los mapas creados es el que refleja el precio medio de las viviendas por distritos, y en el cual observamos que los distritos más caros son Chamberí, Salamanca y Chamartín. Los distritos con los precios más bajos son Arganzuela y Tetuán. En los mapas de esta página podemos observar los datos desagregados y agregados por distritos.

BASE DATOS Idealista

1 Elige un atributo que desees mostrar

PRECIO

+ Agregar atributo

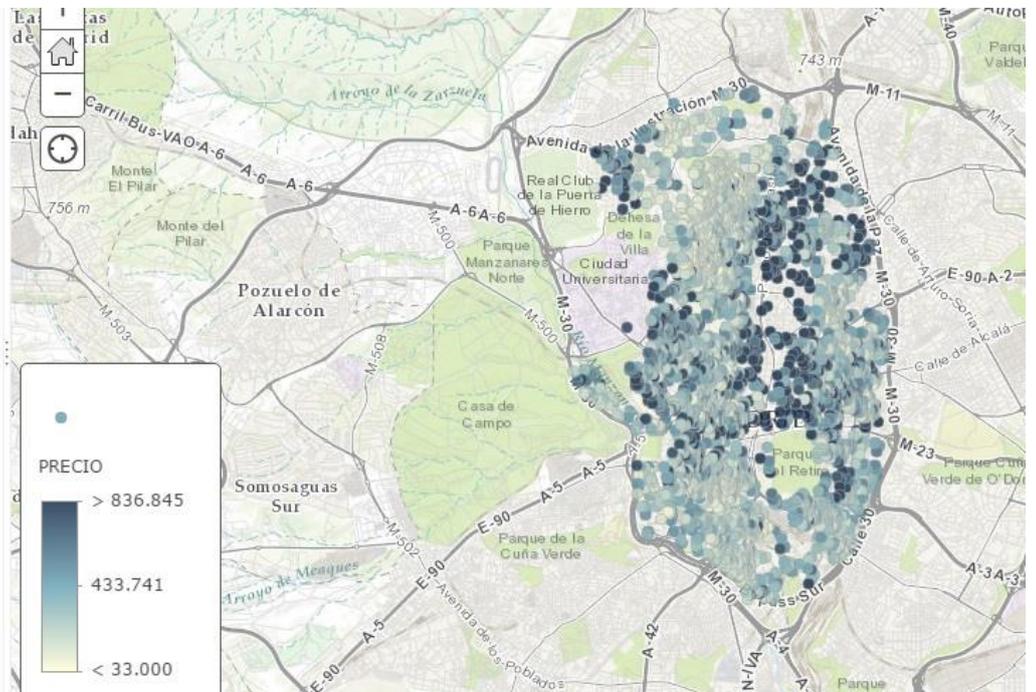
2 Selecciona un estilo de dibujo

Recuentos y cantidades (color)

OPCIONES

Recuentos y cantidades (tamaño)

HECHO CANCELAR



50 Mapa de la variable precio de todas las observaciones de la base de datos de Idealista; fuente: elaboración propia.

BASE DATOS Idealista

1

Elige un atributo que desees mostrar

PRECIO

+ Agregar atributo

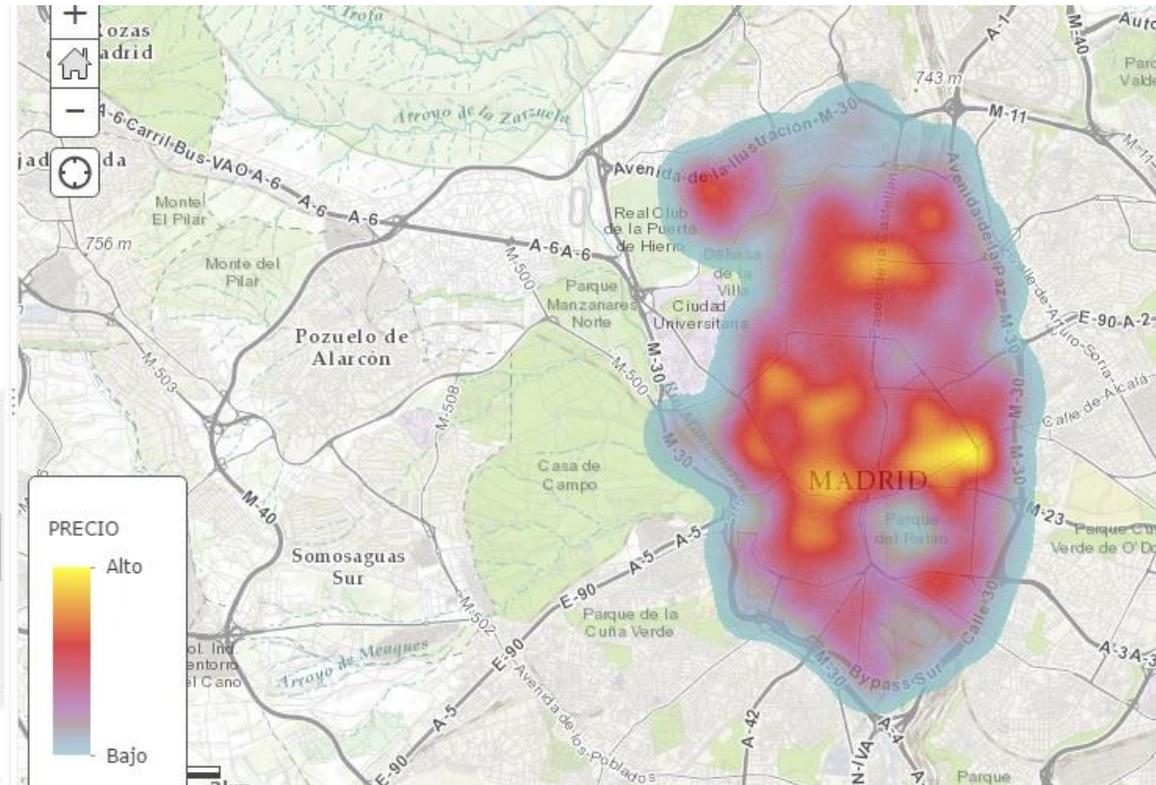
2

Selecciona un estilo de dibujo

Mapa de calor

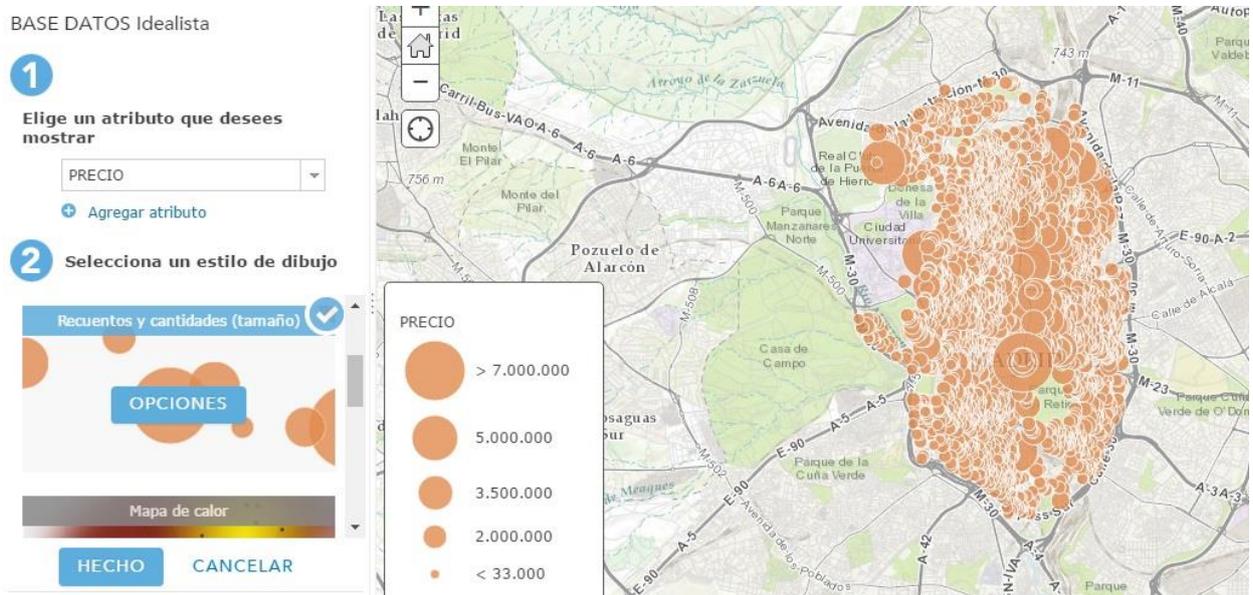
OPCIONES

HECHO CANCELAR

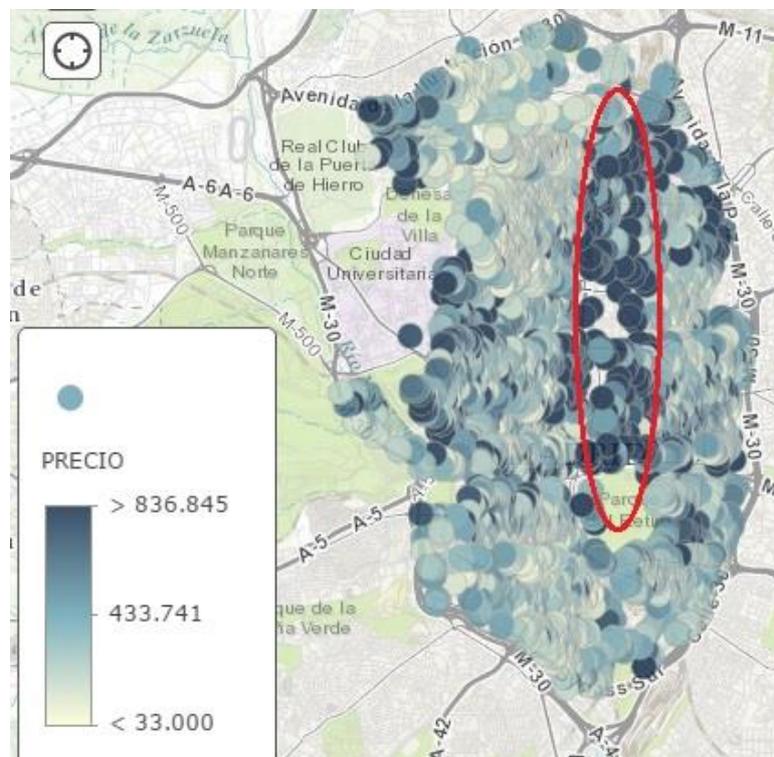


51 Mapa de calor del precio de las viviendas de la base de datos de Idealista; fuente: elaboración propia.

Además del precio de las viviendas en estudio, en los mapas se pueden observar efectos espaciales que guardan coherencia con los resultados obtenidos en los modelos de regresión, como es la constatación del efecto de la cercanía al Paseo de la Castellana (eje norte - sur de la ciudad) en el encarecimiento de los precios.

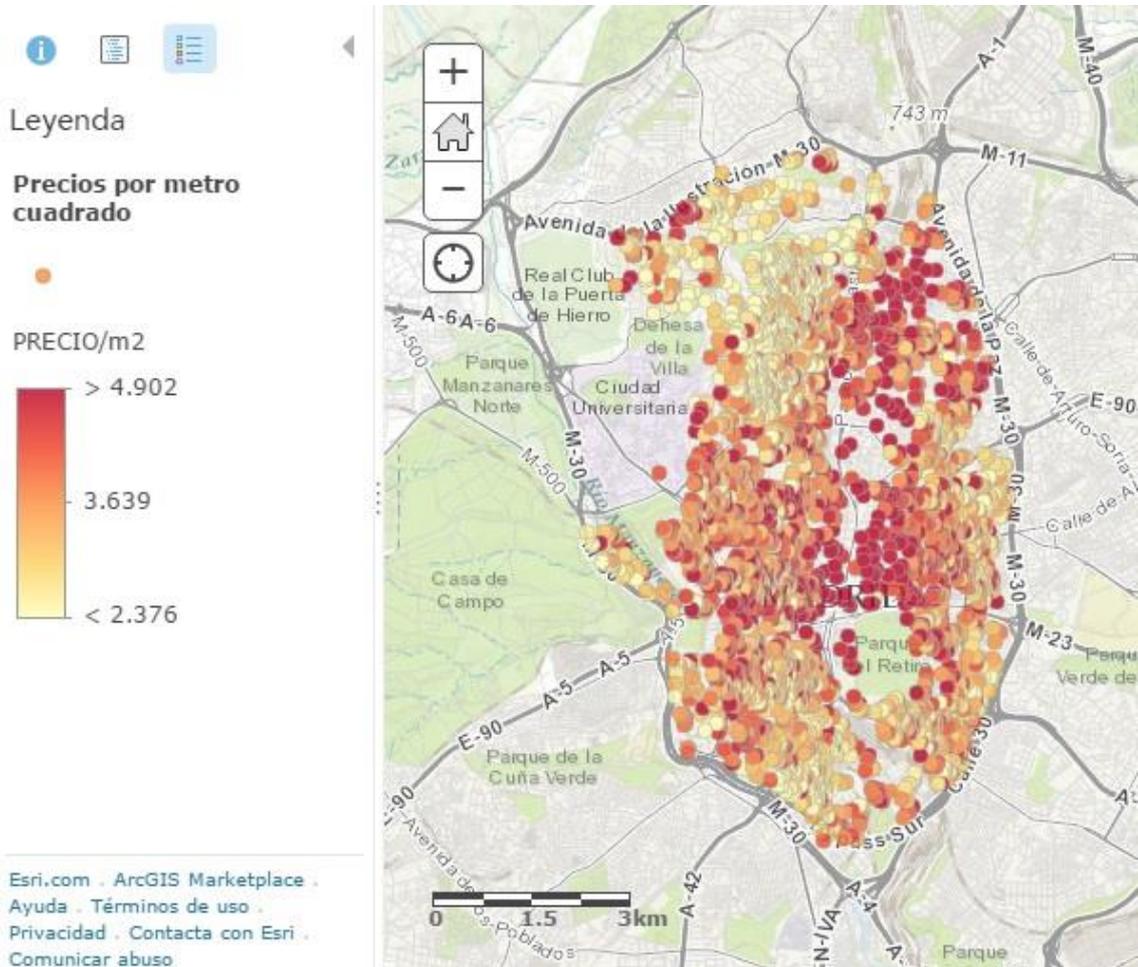


52 Mapa de la variable precio mediante círculos de tamaño; fuente: elaboración propia.



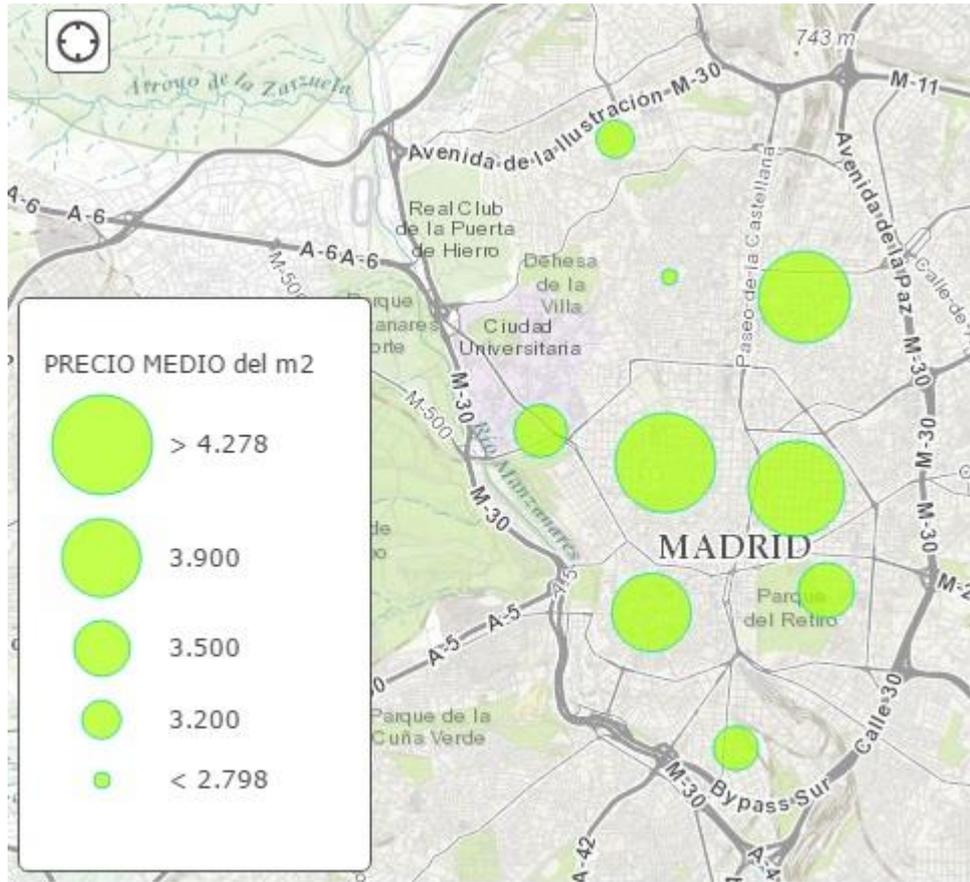
53 Zona de concentración de viviendas de alto precio en torno al Paseo de la Castellana; fuente: elaboración propia.

Como ya hemos dicho en más ocasiones en este trabajo de investigación, en lo referente al precio de la vivienda es siempre interesante analizar el dato del precio por metro cuadrado, pues es el que nos permite comparar zonas acudiendo a un único ratio.

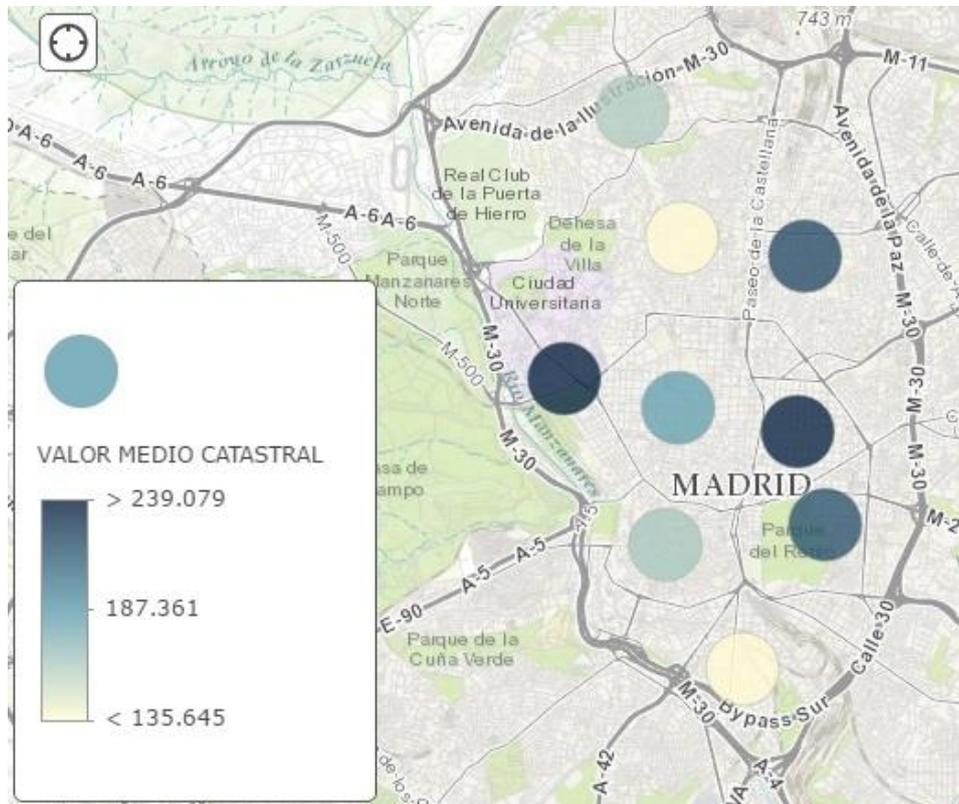


54 Precio por metro cuadrado de toda la base de datos de Idealista; fuente: elaboración propia.

En este caso, y de manera muy similar a la variable *precio*, observamos las zonas de precios más caros por metro cuadrado en los distritos de Salamanca, Chamberí y Chamartín, así como el efecto de la cercanía al Paseo de la Castellana (que por otro lado, tal como indican los resultados de la regresión, es también acorde al efecto de la lejanía a la M30).



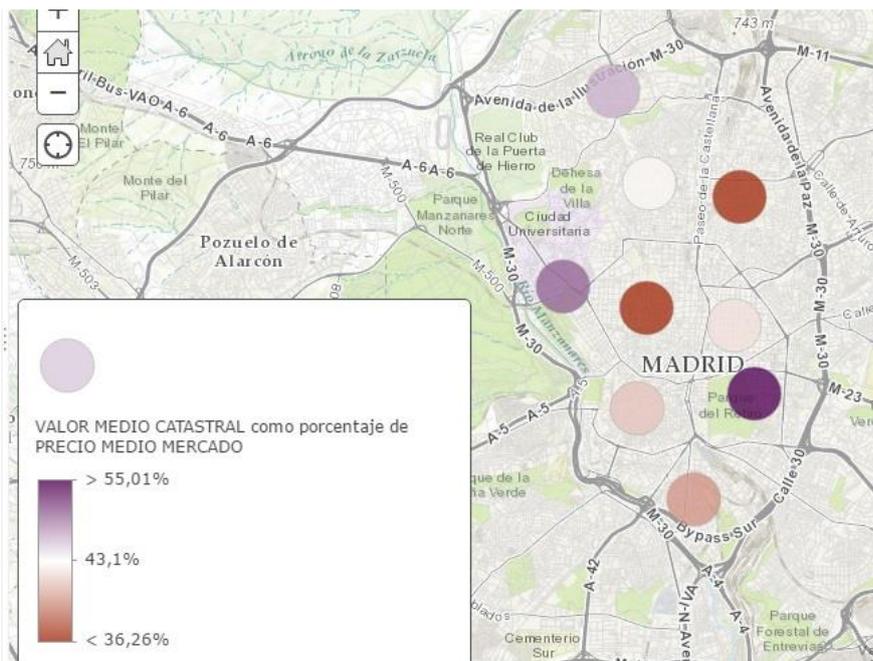
55 Precio medio del metro cuadrado en los distritos analizados; fuente: elaboración propia.



56 Valor catastral medio de la base de datos analizada por distritos; fuente: elaboración propia.

Precio MERCADO Valor CATASTRAL

Muestra VALOR MEDIO CATASTRAL como porcentaje de PRECIO MEDIO MERCADO



57 Porcentaje del valor medio catastral sobre el precio de mercado medio por distritos; fuente: elaboración propia.

Otro de los elementos estudiados en el TFM mediante los mapas de ARCGIS es el valor catastral de las viviendas de la zona estudiada.

En este caso no disponíamos de los valores catastrales de las 4.100 viviendas de la base de datos y sólo pudimos obtener los valores catastrales agregados por distritos. En cualquier caso, al disponer de los precios de mercado medios de cada distrito pudimos establecer relaciones y cuantificar la relación y el ratio entre el valor catastral medio de cada distrito y el precio medio de mercado. Los resultados reflejan que el valor catastral medio en el distrito de Retiro es el más elevado (con valores que rondan el 50% del valor de mercado) mientras que en los distritos de Chamartín y Chamberí los valores catastrales suponen de media alrededor de un 35% de los valores medios de precio de mercado. El estudio pormenorizado de los valores catastrales por zonas y su relación con los precios de mercado es una de las futuras líneas de investigación en el campo de la vivienda y de la econometría espacial que se desprenden del estudio realizado en este trabajo.

## CONCLUSIONES

1. A tenor de los resultados obtenidos en los dos modelos de regresión realizados en el trabajo la primera de las conclusiones es que los factores espaciales sí tienen influencia en la formación de los precios de la vivienda, confirmando lo ya expuesto en otros trabajos similares.
2. El modelo de regresión espacial así como los resultados del índice de Moran y el scatterplot *LISA* reflejan que en la ciudad de Madrid (más concretamente en la zona analizada dentro de la M30) sí existe autocorrelación espacial entre las viviendas respecto de la variable de su precio de mercado. Esto se confirma con el estudio y el análisis de los mapas realizados con *ARCGIS*, donde se comprueba con claridad la existencia de clústeres de viviendas en distintas zonas del espacio estudiado.
3. El precio de la vivienda en la ciudad de Madrid (dentro de la m30) es muy heterogéneo y presenta grandes diferencias entre zonas (distritos y barrios), pero podemos confirmar que el precio medio del metro cuadrado en la ciudad de Madrid en la actualidad se encuentra alrededor de los 3.500 €/m<sup>2</sup> según los resultados del análisis de regresión.
4. Los distritos de Arganzuela y Tetuán presentan los precios más bajos mientras que los distritos de Salamanca, Retiro, Chamberí y Chamartín presentan los precios más altos.
5. Los barrios que presentan precios más altos son Recoletos, El Viso, Castellana y Almagro mientras que los barrios con precios más bajos son Berruguete y Bellas Vistas.
6. Las características de las viviendas que les hacen incrementar su precio son la altura del piso (planta), a razón de unos 10.000 € más por cada planta que se sube, su tamaño, a razón del precio medio por metro cuadrado en el área estudiada, en este caso unos 3.500 €/m<sup>2</sup> y el número de baños a razón de unos 70.000 € por baño adicional. El hecho de que la vivienda sea exterior aumenta el precio en unos 19.000 € atendiendo al coeficiente significativo del *modelo 1*; esto presenta cierta lógica tratándose de la ciudad de Madrid, donde las manzanas de viviendas suelen tener un reducido patio interior y donde se valora mucho el hecho de que una vivienda sea exterior (con balcones y/o ventanas que dan a una calle, y no al patio interior).
7. La variable *weight\_4* confirma que existe una correlación espacial y que las viviendas de alrededor (en este caso el precio de las 4 viviendas más cercanas) de una vivienda estudiada tienen influencia en el precio de la vivienda en cuestión.
8. Del mismo modo, la variable *w\_size* corrobora que, en el caso del tamaño de las viviendas, también existe una correlación espacial, es decir, hay relación entre los tamaños de las viviendas por zonas. Los mapas elaborados con el SIG utilizado en este

TFM confirman estas correlaciones, ya que en ellos se observa de un modo claro la existencia de clústeres de viviendas agrupadas en función de su precio y de su tamaño.

9. Con respecto a las variables con efectos espaciales incorporadas en los dos modelos de regresión realizados, los coeficientes obtenidos nos indican lo siguiente:

- 9.1. En lo relativo a la variable **distancia al km 0 (Puerta del Sol)**, el análisis de regresión nos indica un aumento del precio de las viviendas cuanto más cerca se encuentre la vivienda de este punto, lo que se puede explicar debido a que el Km 0 (donde se encuentra el emblemático reloj de la Puerta del Sol) representa la zona más céntrica de la ciudad de Madrid y además se encuentra en un barrio de precios altos con respecto a la media (distrito centro). La cercanía a cascos urbanos y zonas céntricas casi siempre supone un aumento de los precios de la vivienda. Además, esto se puede observar en el hecho de que los distritos más caros de la ciudad son Centro, Salamanca y Chamberí, los más cercanos a la Puerta del Sol.

- 9.2. En cuanto a la variable **distancia al Paseo de la Castellana (eje norte-sur)**, el análisis de regresión indica un aumento del precio de las viviendas cuanto más cerca se encuentren las viviendas de esta vía, si bien el coeficiente significativo lo es para la franja comprendida entre los 300 m y los 700 m de distancia al eje. Este resultado es coherente con los resultados obtenidos en los mapas de georreferenciación realizados con ARCGIS, donde se observa con claridad que las viviendas cercanas a la Castellana son las que presentan los precios más elevados de la ciudad en términos de precio del metro cuadrado. Conviene recordar que en los análisis de regresión realizados en este trabajo la variable dependiente es el precio total de las viviendas, pero en el estudio del mercado inmobiliario de una determinada zona o ciudad, el indicador más significativo es el precio del metro cuadrado. Además, en este caso se observa que los precios más altos se dan en barrios como Almagro, El Viso o Castellana, barrios que se articulan a lo largo del eje que configura esta vía.

- 9.3. En cuanto a la variable **distancia a la M30**, el análisis de regresión indica un aumento del precio de las viviendas cuanto más lejos se encuentren las viviendas de esta vía de circunvalación, si bien el coeficiente significativo lo es para la franja comprendida entre los 100 m y los 300 m de distancia a la vía. En este caso, los motivos de esta disminución del precio en las cercanías a dicha vía se explican principalmente por los efectos del ruido, ya que es una vía muy transitada constantemente y a todas horas del día y vivir cerca de ella reduce el nivel de confort asociado al descanso. Este resultado guarda coherencia con el coeficiente de distancia al Paseo de la Castellana y además se ve bien reflejado en los mapas realizados en el trabajo de investigación. Del mismo modo que en los mapas se observa que las zonas con los valores de precio/m<sup>2</sup> más elevados se encuentran en las zonas aledañas al Paseo de la Castellana, los valores de precio/m<sup>2</sup> más bajos se dan en zonas cercanas a la vía de circunvalación de la M30. Además, los precios de la vivienda en la ciudad de Madrid difieren mucho entre las viviendas situadas

dentro del anillo perimetral que supone esta vía y las que se encuentran fuera, pudiéndose llegar a considerar a la M30 una barrera o límite claro en el estudio de los precios de las viviendas en la ciudad de Madrid.

**9.4.** En lo relativo a la variable **distancia a los nudos de la M30**, mientras que en el *modelo 1* el coeficiente nos indica un aumento del precio cuanto más cerca se encuentre la vivienda de dichos nudos, la regresión con las distancias segmentadas del *modelo 2* nos indica que en el área comprendida entre los 400 y los 800 metros de radio de la M30 el precio disminuye mientras que el coeficiente relativo a las viviendas que se encuentran a menos de 400 m no es significativo. Ambos resultados, que en principio pueden parecer contradictorios, reflejan la dificultad que existe en explicar el precio de la vivienda con factores espaciales como éste, ya que la cercanía a un nudo de comunicación de la M30 puede implicar a la vez facilidad de acceso a vías de salida y entrada a la ciudad tanto como la cercanía a una zona ruidosa y de mucho tráfico. Por ello, con los resultados relativos a esta variable no se puede determinar con precisión una correlación espacial, ya que los resultados de ambos modelos de regresión no nos otorgan evidencia suficiente de que ésta exista.

**9.5.** En lo relativo a la variable **distancia a las estaciones de metro**, el hecho de que no sea significativo para la formación de la variable “precio” (el coeficiente no es significativo en el *modelo 1*) se puede deducir del hecho de que, dentro de la M30, en la ciudad de Madrid hay numerosas estaciones de metro y prácticamente ninguna zona se encuentra vacía de paradas de metro, con lo cual no es un factor especialmente explicativo en la formación del precio de las viviendas en el área estudiada en este trabajo. Además, podemos intuir que el vivir cerca de una parada de metro debería de resultar en un aumento del precio del inmueble por la comodidad de encontrarse cerca del acceso a la red de transporte público y esto es coherente con el resultado obtenido en el *modelo 2* de regresión, donde el coeficiente significativo nos indica que las viviendas situadas a menos de 50 metros sí presentan un alza en el precio. Esto indica que las viviendas muy cercanas a una de ellas (a menos de 50 metros) sí presentan un precio superior explicado por esta variable.

**9.6.** En cuanto a la variable **distancia a las zonas verdes (parques y jardines)**, el coeficiente del *modelo 1* nos indica el aumento del precio cuanto más cerca se encuentre la vivienda de una zona verde, más concretamente a razón de unos 60.000 € por cada Km. En el *modelo 2*, el modelo de regresión con las distancias segmentadas, vemos que las viviendas situadas a menos de 100 metros de una zona verde ven incrementado su precio, algo que también ocurre con las situadas a menos de 300 m de un parque o jardín. En esta variable se observa claramente cómo el precio de las viviendas no presenta relación lineal alguna con las variables espaciales (distancias), ya que, según los coeficientes del *modelo 2*, una vivienda situada a menos de 100 m de una zona verde incrementaría su precio

aproximadamente en unos 58.000 € mientras que las situadas en la franja entre los 100 y los 300 m lo harían en unos 18.000 €. Por ello podríamos concluir que son las viviendas más cercanas a los parques (a menos de 100 m) las que ven notablemente incrementado su valor de mercado. Esto presenta cierta lógica por el hecho de que habitar en una ciudad excesivamente contaminada y ruidosa como Madrid en las cercanías de un parque o jardín incrementa la calidad de vida en cuanto a mejor calidad del aire respirado.

**9.7.** En cuanto a la variable **distancia a hospitales**, el signo negativo del coeficiente en el *modelo 1* nos indica el aumento de precio del inmueble por cada Km de cercanía a un hospital. En el *modelo 2* observamos de nuevo que la relación espacial no es lineal y que, para las zonas cercanas a un hospital a menos de 300 m, las viviendas experimentan subidas de precio del orden de los 40.000 €, mientras que para las ubicadas en la franja comprendida entre los 300 y los 600 m el precio se ve incrementado alrededor de los 26.000 €. Ambos resultados confirman que la cercanía a un hospital es uno de los factores que incrementa el precio de la vivienda en la zona estudiada y nos ayuda a entender la importancia que se da a la hora de elegir el lugar de residencia a la cercanía a un hospital. También es interesante el hecho de que el valor del coeficiente de *distancia a zonas verdes* prácticamente triplica al de cercanía a un hospital. Hay que tener en cuenta que en este trabajo se han referenciado solo los grandes hospitales de referencia de la ciudad de Madrid, y no los centros de salud.

**9.8.** En cuanto a la variable **distancia a colegios**, el coeficiente del *modelo 1* nos indica la disminución del precio cuanto más cerca se encuentre la vivienda de un colegio, mientras que los coeficientes en el *modelo 2* con las distancias segmentadas nos indican que las viviendas situadas a menos de 100 m y a menos de 300 m de un colegio disminuyen su precio. Ambos resultados confirman que la cercanía a un colegio es uno de los factores que disminuye el precio de la vivienda en la zona estudiada. Esto puede deberse al ruido y al bullicio existente en los alrededores de un colegio, y también cabe resaltar que en este trabajo se han tenido en cuenta solamente los colegios públicos de la zona estudiada, sin tener en cuenta en el modelo los colegios concertados ni los privados, que pudieran arrojar resultados diferentes.

**9.9.** En cuanto a la variable **distancia a museos** el coeficiente del *modelo 1* nos indica la disminución del precio cuanto más cerca se encuentre la vivienda de un museo, mientras que los coeficientes en el *modelo 2* con las distancias segmentadas no son significativos. Este resultado indica que la cercanía a un museo disminuye el precio de la vivienda en la zona estudiada.

En resumen, y a modo de conclusión general, en lo relativo a las distancias estudiadas en este trabajo, que representan la influencia de factores espaciales como factores explicativos de los precios de la vivienda, podemos concluir que hay una influencia al alza en los precios de mercado de las viviendas que se encuentran cerca de zonas verdes (parques y jardines), hospitales, y también en aquéllas situadas cerca del Paseo de la Castellana, de la Puerta del Sol y muy cercanas a una estación de metro (a menos de 50 metros).

Por el contrario, hay una influencia a la baja en los precios de mercado de las viviendas que se encuentran cerca de la M30, de museos y de colegios (públicos).

Por último, el factor de distancia a los nudos de la M30 no puede ser explicado con claridad mediante los modelos utilizados en este trabajo, ya que los resultados obtenidos en ambos modelos son contradictorios.

## INDICE DE GRÁFICOS, TABLAS E IMÁGENES

1 Variación interanual de los precios de la vivienda en la zona costera de la provincia de Alicante y del Mar Menor; fuente: CartoDB y OpenStreetMap. ....	6
2 Evolución del precio de la vivienda por metro cuadrado; fuente: web El Economista. ....	7
3 Variación interanual del precio de la vivienda; fuente: web EL PAIS con datos del INE. ....	7
4 Precio medio de la vivienda en Madrid; fuente: web EL PAIS. ....	8
5 Variación anual por distritos en Madrid; fuente: web EL PAIS. ....	9
6 Reloj inmobiliario; fuente: La Caixa Research con datos del INE ....	10
7 Índice de accesibilidad para la compra de vivienda; fuente: Sociedad de Tasación. ....	11
8 Variación interanual del precio de la vivienda por Comunidades Autónomas; fuente: TINSAs ....	12
9 Compraventa de viviendas 2010 – 2016; fuente: Instituto Nacional de Estadística ....	12
10 Características de la vivienda deseada por los españoles; fuente: elmundo.es.....	15
11 Noticia del 11 de Septiembre de 2016 sobre la influencia sobre el precio de la vivienda de la cercanía a colegios en la ciudad de Madrid; fuente: elmundo.es ....	21
12 Variación interanual del precio de la vivienda en la costa española; fuente: OpenStreetMap y CartoDB con datos de TINSAs.....	23
13 Variación interanual del precio de la vivienda en Cartagena; fuente: OpenStreetMap y CartoDB con datos de TINSAs. ....	23
14 Cálculo de una ruta óptima para vehículos entre un punto de origen (en verde) y un punto de destino (en rojo) a partir de datos del proyecto OpenStreetMap. ....	24
15 Un SIG puede mostrar la información en capas temáticas para realizar análisis multicriterio complejos. ....	25
16 Un ejemplo de uso de la superposición de capas en una aplicación SIG ....	26
17 Resultados elecciones generales en el municipio de Madrid; fuente: El Confidencial.LAB ....	27
18 Mapa de representación de las multas de tráfico en Madrid dentro de la M30; fuente: ElPaís.com...	27
19 Análisis de datos espaciales de la ciudad de Londres; fuente: OpenStreetMap.org.....	28
20 Mapa sobre potenciales compradores en zonas comerciales; fuente: web de CartoDB. ....	29
21 Georreferenciación de los meteoritos caídos en la Tierra y su tamaño; fuente: CartoDB.....	30
22 Mapa sobre precios de alquileres de viviendas de Airbnb; fuente: web de CartoDB. ....	31
23 Aplicación para Smartphone de búsqueda de viviendas mediante georreferenciación; fuente: Idealista.com. ....	32
24 Georreferenciación de zonas verdes en la zona analizada del municipio de Madrid; fuente: elaboración propia. ....	32
25 Elementos de una ecuación de regresión de OLS; fuente: tutorial de regresiones de ArcGIS.....	33
26 Técnicas empleadas en el AEDE; fuente: Luc Anselin (1988) ....	37
27 Informe con datos sobre valoración de una vivienda en Madrid; fuente: web de precioviviendas. ....	41
28 Evolución del precio por metro cuadrado en la ciudad de Madrid y en la Región de Murcia; fuente: web de TINSAs. ....	42
29 Datos sobre el precio de las viviendas en la ciudad de Madrid; fuente: web de TINSAs.....	43
30 Tabla de parámetros para la obtención de datos; fuente: elaboración propia a partir de Idealista.....	47
31 Captura de información para la extracción de la base de datos; fuente: web de Idealista. ....	47
32 Área de análisis y extracción de la base de datos del TFM dentro de la M30; fuente: Idealista.com...	48
33 Distritos del municipio de Madrid; fuente: munimadrid.org ....	48
34 Web de Datos Abiertos del Ayuntamiento de Madrid; fuente: web de datos.madrid.es. ....	49
35 Georreferenciación de elementos en Maps y en ARCGIS online; fuente: elaboración propia.....	50
36 Georreferenciación de colegios y estaciones de metro; fuente: elaboración propia. ....	50
37 Tamaño medio de las viviendas por distritos; fuente: elaboración propia. ....	51
38 Número de habitaciones de las viviendas por distritos; fuente: elaboración propia.....	52

39 Precio medio de las viviendas por distritos; fuente: elaboración propia. ....	52
40 Tamaño medio de las viviendas analizadas por distritos; fuente: elaboración propia. ....	53
41 Tabla de tamaños y precios por distritos; fuente: elaboración propia. ....	53
42 Barrios con precio medio de mercado más alto; fuente: elaboración propia. ....	54
43 Precio medio de mercado de todos los barrios analizados en el trabajo; fuente: elaboración propia. ....	55
44 Índice de Moran; fuente: elaboración propia con el software GeoDa. ....	57
45 LISA significance map; fuente: elaboración propia con el software GeoDa. ....	58
46 LISA cluster map; fuente: elaboración propia con el software GeoDa. ....	59
47 Tabla de resultados de los modelos de regresión; fuente: elaboración propia. ....	61
48 Base de datos por distritos utilizada en el TFM; fuente: elaboración propia. ....	67
49 Precio medio de mercado de la base de datos analizada por distritos; fuente: elaboración propia. ....	68
50 Mapa de la variable precio de todas las observaciones de la base de datos de Idealista; fuente: elaboración propia. ....	68
51 Mapa de calor del precio de las viviendas de la base de datos de Idealista; fuente: elaboración propia. ....	69
52 Mapa de la variable precio mediante círculos de tamaño; fuente: elaboración propia. ....	70
53 Zona de concentración de viviendas de alto precio en torno al Paseo de la Castellana; fuente: elaboración propia. ....	70
54 Precio por metro cuadrado de toda la base de datos de Idealista; fuente: elaboración propia. ....	71
55 Precio medio del metro cuadrado en los distritos analizados; fuente: elaboración propia. ....	72
56 Valor catastral medio de la base de datos analizada por distritos; fuente: elaboración propia. ....	72
57 Porcentaje del valor medio catastral sobre el precio de mercado medio por distritos; fuente: elaboración propia. ....	73
58 Precio de mercado y valor catastral medio de los barrios del distrito Centro; fuente: elaboración propia. ....	85
59 Precio de mercado y valor catastral medio de los barrios del distrito Arganzuela; fuente: elaboración propia. ....	85
60 Precio de mercado y valor catastral medio de los barrios del distrito Retiro; fuente: elaboración propia. ....	86
61 Precio de mercado y valor catastral medio de los barrios del distrito Salamanca; fuente: elaboración propia. ....	86
62 Precio de mercado y valor catastral medio de los barrios del distrito Chamartín; fuente: elaboración propia. ....	87
63 Precio de mercado y valor catastral medio de los barrios del distrito Tetuán; fuente: elaboración propia. ....	87
64 Precio de mercado y valor catastral medio de los barrios del distrito Chamberí; fuente: elaboración propia. ....	88
65 Precio de mercado y valor catastral medio de los barrios del distrito Fuencarral - El Pardo; fuente: elaboración propia. ....	88
66 Precio de mercado y valor catastral medio de los barrios del distrito Moncloa - Aravaca; fuente: elaboración propia. ....	89
67 Georreferenciación de la Puerta del Sol; fuente: elaboración propia con ARCGIS online. ....	90
68 Georreferenciación del Paseo de la Castellana; fuente: elaboración propia con ARCGIS online. ....	90
69 Georreferenciación de la M30; fuente: elaboración propia con ARCGIS online. ....	91
70 Georreferenciación de los nudos de la M30; fuente: elaboración propia con ARCGIS online. ....	91
71 Georreferenciación de las estaciones de metro; fuente: elaboración propia con ARCGIS online. ....	91
72 Georreferenciación de los colegios públicos; fuente: elaboración propia con ARCGIS online. ....	91
73 Georreferenciación de las zonas verdes (parques y jardines); fuente: elaboración propia con ARCGIS online. ....	91
74 Georreferenciación de los hospitales; fuente: elaboración propia con ARCGIS online. ....	91

75 Georreferenciación de los museos; fuente: elaboración propia con ARCGIS online. ....91

## BIBLIOGRAFÍA

ANSELIN, LUC. **Local indicators of spatial association - LISA.** *Geographical analysis*, 27, 93-115. 1995.

ANSELIN, LUC. **Visualizing multivariate spatial correlation with dynamically linked windows.** *New tools for spatial data analysis, proceedings of a workshop.* 2002.

ANSELIN, LUC. **Exploring spatial data with Geoda: A Workbook.** *Spatial Analysis Laboratory, Department of Geography, University of Illinois.* 2005.

ANSELIN, LUC. **Geoda: an introduction to spatial data analysis.** *Geographical analysis*, 38, 5-22. 2005.

BASU, SABYASACHI, THIBODEAU, THOMAS G. **Analysis of spatial autocorrelation in house prices.** *The Journal of Real Estate Finance and Economics*, 17, 61-85. 1998.

CAN, AYSE. **The measurement of neighborhood dynamics in urban house prices.** *Economic Geography*, 66, 254-272. 1990.

CAN, AYSE; MEGBOLUGBE, ISAAC. **Spatial dependence and house price index construction.** *The Journal of Real Estate Finance and Economics*, 14, 203-222. 1997.

CHASCO YRIGOYEN, CORO. **Econometría espacial aplicada a la predicción-extrapolación de datos microterritoriales.** Tesis Doctoral. Comunidad de Madrid; Consejería de Economía e Innovación Tecnológica; Dirección General de Economía y Planificación. Madrid, 2003.

CHASCO YRIGOYEN, CORO; LE GALLO, JULIE. **Spatial analysis of urban growth in Spain, 1900–2001.** *Empirical Economics*, 34, 59-80. 2008.

CHASCO YRIGOYEN, CORO; SÁNCHEZ REYES, BEATRIZ. **Externalidades ambientales y precio de la vivienda en Madrid: un análisis con regresión cuantílica espacial.** *Revista Galega de Economía*, 21, 2. 2012.

CHASCO YRIGOYEN, CORO; LE GALLO, JULIE. **The impact of objective and subjective measures of air quality and noise on house prices: a multilevel approach for downtown Madrid.** *Economic Geography*, 89, 127-148. 2013.

GÁMEZ MARTÍNEZ, MATÍAS. **Nuevas técnicas de estadística espacial para la economía: modelización del precio de la vivienda libre en la ciudad de Albacete.** Tesis doctoral, Universidad de Castilla La Mancha. 1998.

GÁMEZ MARTÍNEZ, MATÍAS; MONTERO LORENZO, JOSÉ MARÍA; GARCÍA RUBIO, NOELIA. **Kriging methodology for regional economic analysis: estimating the housing price in Albacete.** *International advances in Economic Research*, 6, 438-451. 2000.

GÁMEZ MARTÍNEZ, MATÍAS; GARCÍA RUBIO, NOELIA; ALFARO CORTÉ, ESTEBAN. **Una aplicación de la estadística espacial al comportamiento de la vivienda de alquiler en España.** Comunicación presentada en la XVIII Reunión ASEPELT (*Asociación Internacional de Economía Aplicada*)-España. León. 2004.

GAO, XIAOLU; ASAMI, YASUSHI. **Influence of spatial features on land and housing prices.** *Tsinghua Science & Technology*, 10, 344-353. 2005.

GARCÍA MONTALVO, JOSÉ. **Un análisis empírico del crecimiento del precio de la vivienda en las comunidades autónomas españolas.** *Revista Valenciana de Economía y Hacienda*, 2, 117-136. 2001.

GENG JIJIN; CAO KAI; YU LE; TANG YONG. **Geographically weighted regression model (GWR) based spatial analysis of house price in Shenzhen.** 19<sup>th</sup> *International Conference on Geoinformatics*. 2011.

GETIS, ARTHUR. **The analysis of spatial association by use of distance statistics.** *Geographical Analysis* 24, 3. 1992.

GOODCHILD, MICHAEL F. **Spatial Autocorrelation.** *Catmog* 47, *Geo Books*. 1986.

GRIFFITH, DANIEL. **Spatial Autocorrelation: A Primer.** *Resource Publications in Geography, Association of American Geographers*. 1987.

HEEHO, KIM; SAE WOON, PARK; SUNHAE, LEE; XINGQUN, XUE. **Determinants of house prices in Seoul: a quantile regression approach.** *Pacific Rim Property Research Journal*, 21, 91-113, 2015.

LÓPEZ HERNÁNDEZ, FERNANDO ANTONIO; PALACIOS SÁNCHEZ, M<sup>a</sup> ÁNGELES; RUIZ MARÍN, MANUEL. **El desempleo en términos de localización. Una aplicación a las provincias españolas.** *Anales de Economía Aplicada*. Comunicación presentada en la XV Reunión ASEPELT (*Asociación Internacional de Economía Aplicada*)-España. La Coruña. 2001.

LÓPEZ HERNÁNDEZ, FERNANDO ANTONIO; CHASCO YRIGOYEN, CORO; LE GALLO, JULIE. **Exploring scan methods to test spatial structure with an application to housing prices in Madrid.** *Papers in Regional Science*, 94, 229-441. 2015.

MITCHELL, ANDY. **La Guía de ESRI para el análisis SIG**, volumen 2. *ESRI Press*, 2005.

MONTERO LORENZO, JOSÉ MARÍA. **El precio medio del metro cuadrado de la vivienda libre: una aproximación metodológica desde la perspectiva de la geoestadística.** *Estudios de economía aplicada*, 22, 675-693. 2004.

MONTERO LORENZO, JOSÉ MARÍA; LARRAZ IRIBAS, BEATRIZ. **Estimación espacial del precio de la vivienda mediante métodos de krigeado,** *Estadística Española*, 48, 162, 201-240. 2006.

MONTERO LORENZO, JOSÉ MARÍA; FERNÁNDEZ-AVILÉS, GEMA; MÍNGUEZ, ROMÁN. **Spatial hedonic housing price models accounting for spatial autocorrelation, spatial heterogeneity and non-linearity. Prediction and assessment of environment impacts.** Presented at the 25<sup>th</sup> annual *TIES Conference*, Al Ain University, United Arab Emirates. 2015.

MONTERO LORENZO, JOSÉ MARÍA; FERNÁNDEZ AVILÉS, GEMA. **La importancia de los efectos espaciales en la predicción del precio de la vivienda. Una aplicación geoestadística en España.** Departamento de Economía y Empresa. Facultad de Ciencias Jurídicas y Sociales de Toledo. Universidad de Castilla-La Mancha.

MORAN, PATRICK ALFRED PIERCE. **The interpretation of statistical maps.** *Journal of the Royal Statistical Society*, 10, 243-251. 1948.

MORENO JIMÉNEZ, ANTONIO. **Análisis de la situación y evolución del reequilibrio territorial en la ciudad de Madrid.** Universidad Autónoma de Madrid. Ayuntamiento de Madrid, Área de gobierno de economía y participación ciudadana. Madrid, 2006.

MORILLO, MARÍA DEL CARMEN; GARCÍA CEPEDA, FRANCISCO; MARTÍNEZ-CUEVAS, SANDRA. **The application of spatial analysis to cadastral zoning of urban areas: an example in the city of Madrid.** *Survey Review*. 2016.

NAVARRETE ÁLVAREZ, MÓNICA. **Modelos geoestadísticos del precio de la vivienda: aproximación al conocimiento intraurbano de la ciudad de Madrid.** Tesis Doctoral. Universidad Autónoma de Madrid; Facultad de Ciencias Económicas y Empresariales. Madrid, 2011.

PALACIOS SÁNCHEZ, M<sup>a</sup> ÁNGELES; LÓPEZ HERNÁNDEZ, FERNANDO ANTONIO. **Distintos modelos de dependencia espacial. Análisis de autocorrelación.** *Anales de Economía Aplicada*. Comunicación presentada en la XIV Reunión ASEPELT (*Asociación Internacional de Economía Aplicada*)-España. Oviedo. 2000.

POSADA, L. J. **Fundamentos económicos-espaciales de la teoría de Centros de Desarrollo.** *Agricultura y Sociedad*, Ministerio de Agricultura y Comercio de España, 6, 137-180. 1978.

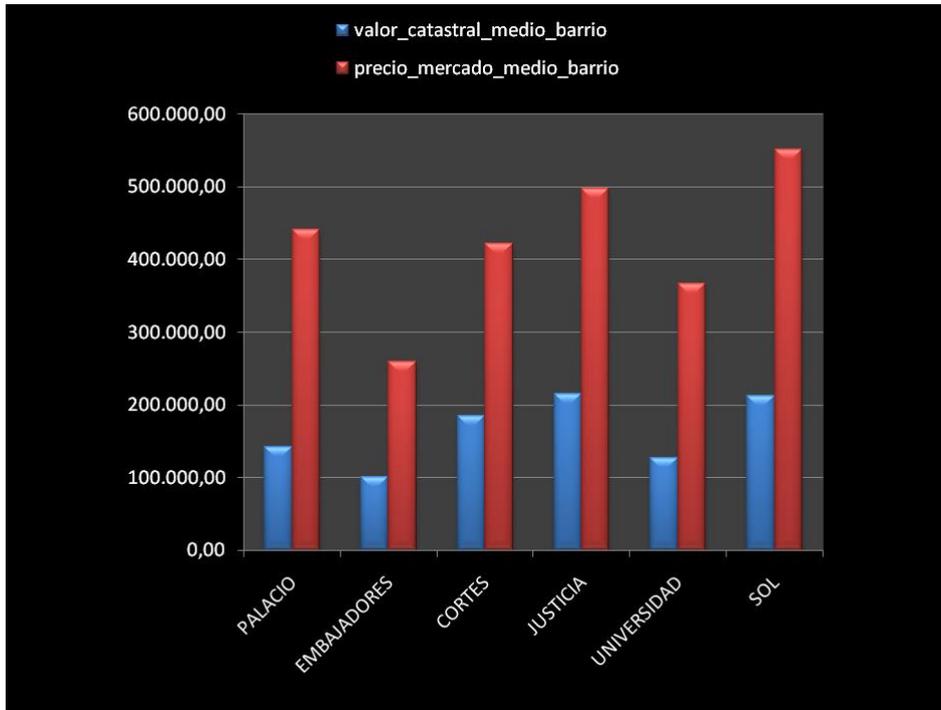
QINGYUAN, TANG; WEI, XU; FULI, AI. **Spatial pattern and structural determinants of Shanghai's housing price: a GWR-based approach.** 19<sup>th</sup> *International Conference on Geoinformatics*. 2011.

VAYÁ VALCARCE, ESTHER; MORENO SERRANO, ROSINA. **Técnicas econométricas en el tratamiento de datos espaciales: la econometría espacial.** *Edicions Universitat Barcelona*. 2000.

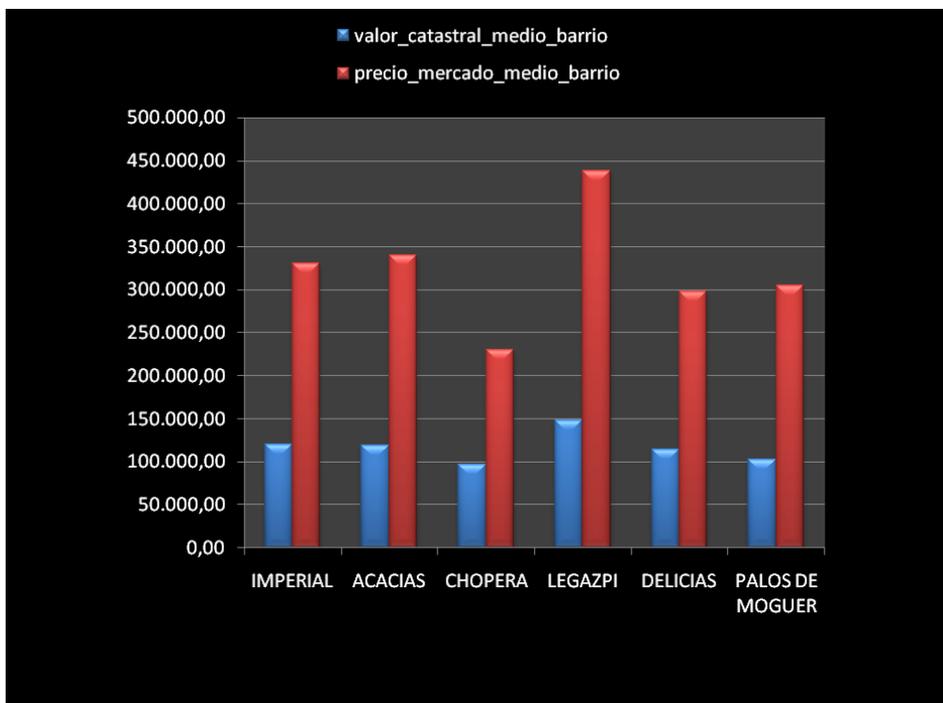
ZHAOHUI, HU. **Spatial econometric analysis of housing price of chinese provinces.** 2<sup>nd</sup> *International Conference on Business Computing and Global Informatization*. 2012.

ZIETZ, JOACHIM; ZIETZ, EMILY NORMAN; SIRMANS, G. STACY. **Determinants of house prices: a quantile regression approach.** *The Journal of Real Estate Finance and Economics*, 37, 317-333. 2008.

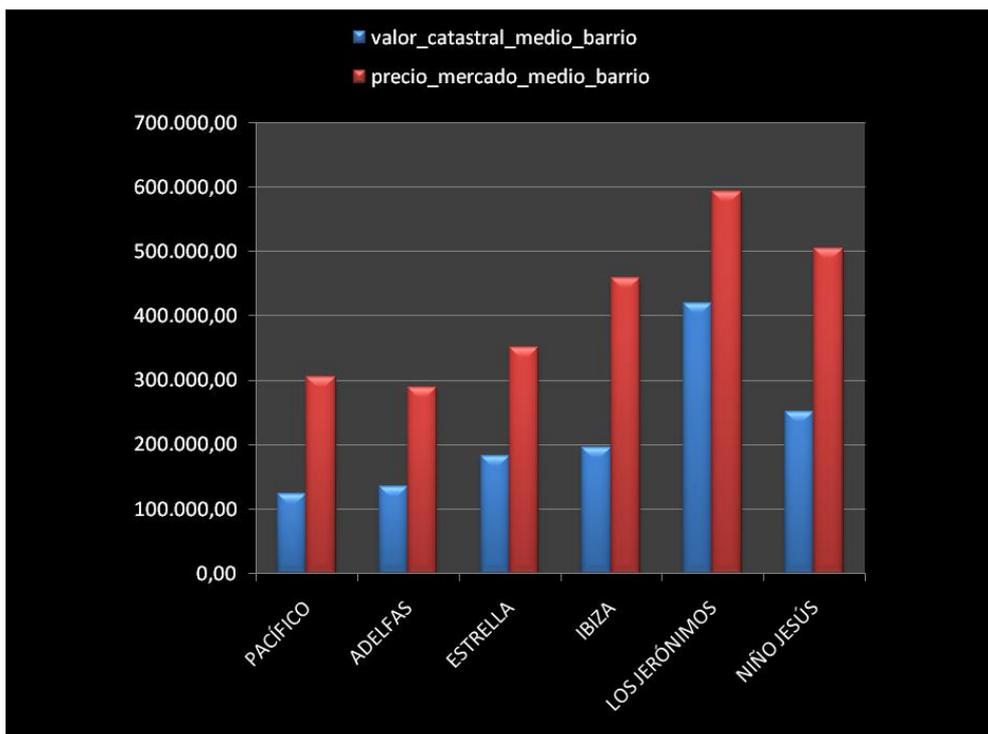
### Anexo 1: precio de mercado y valor catastral por barrios



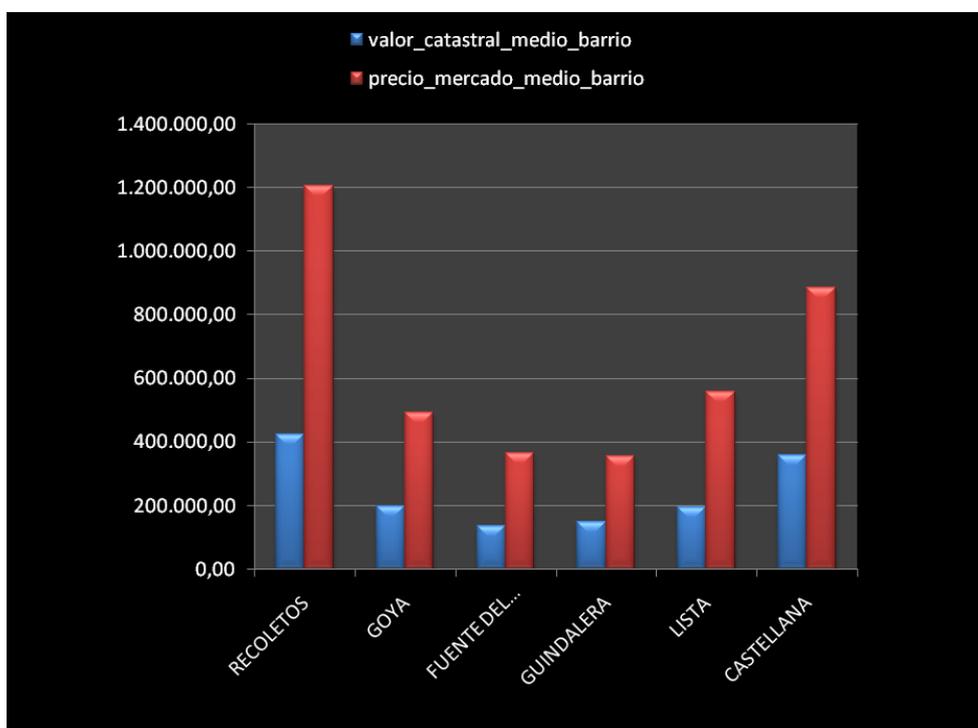
58 Precio de mercado y valor catastral medio de los barrios del distrito Centro; fuente: elaboración propia.



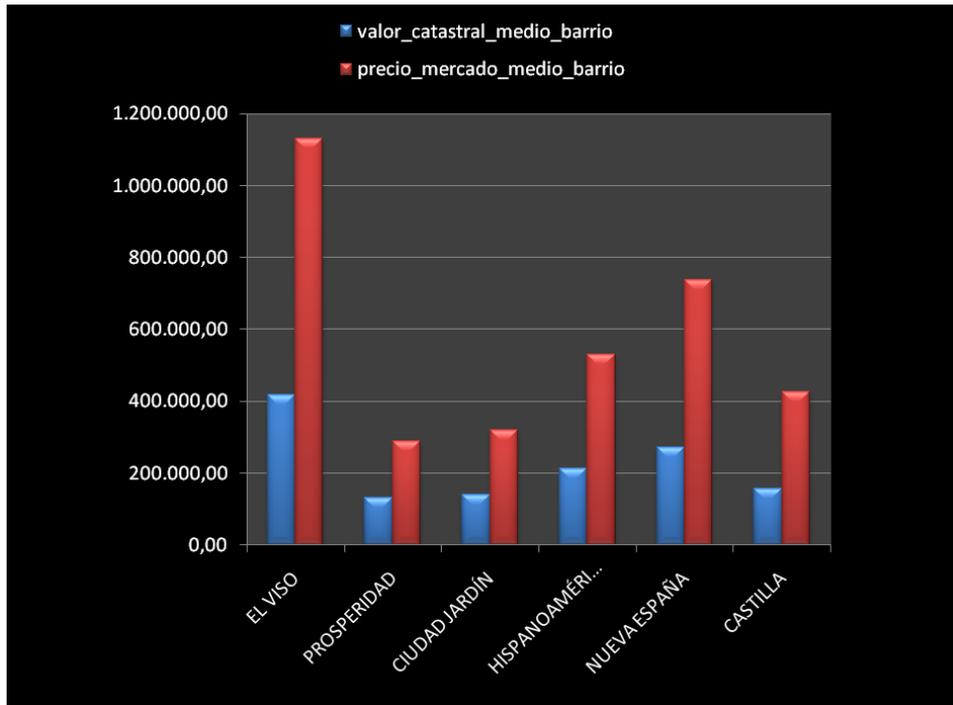
59 Precio de mercado y valor catastral medio de los barrios del distrito Arganzuela; fuente: elaboración propia.



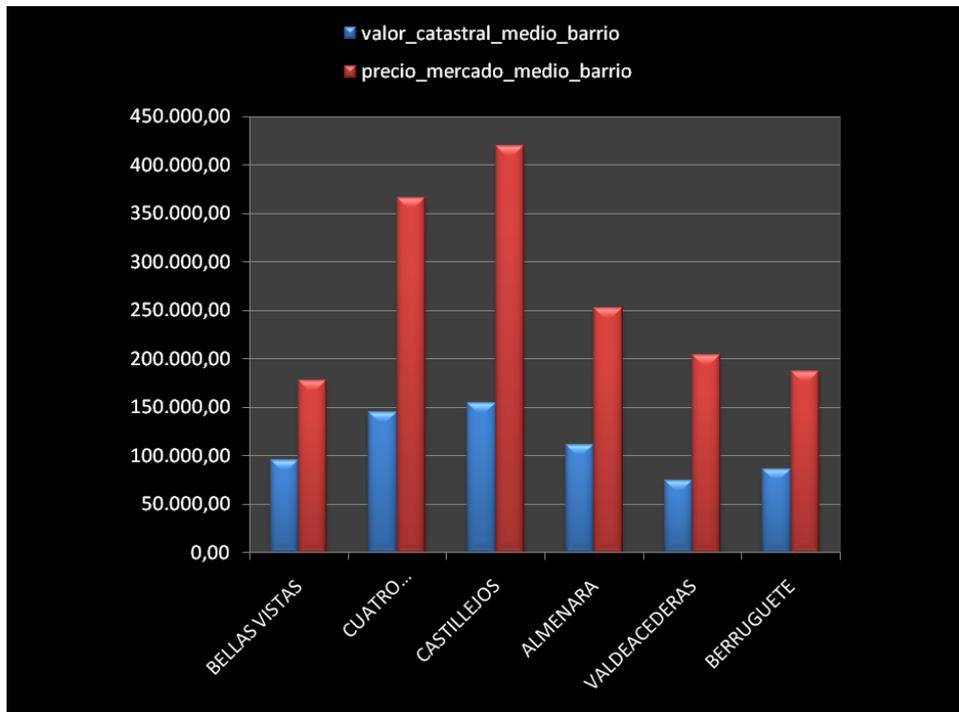
60 Precio de mercado y valor catastral medio de los barrios del distrito Retiro; fuente: elaboración propia.



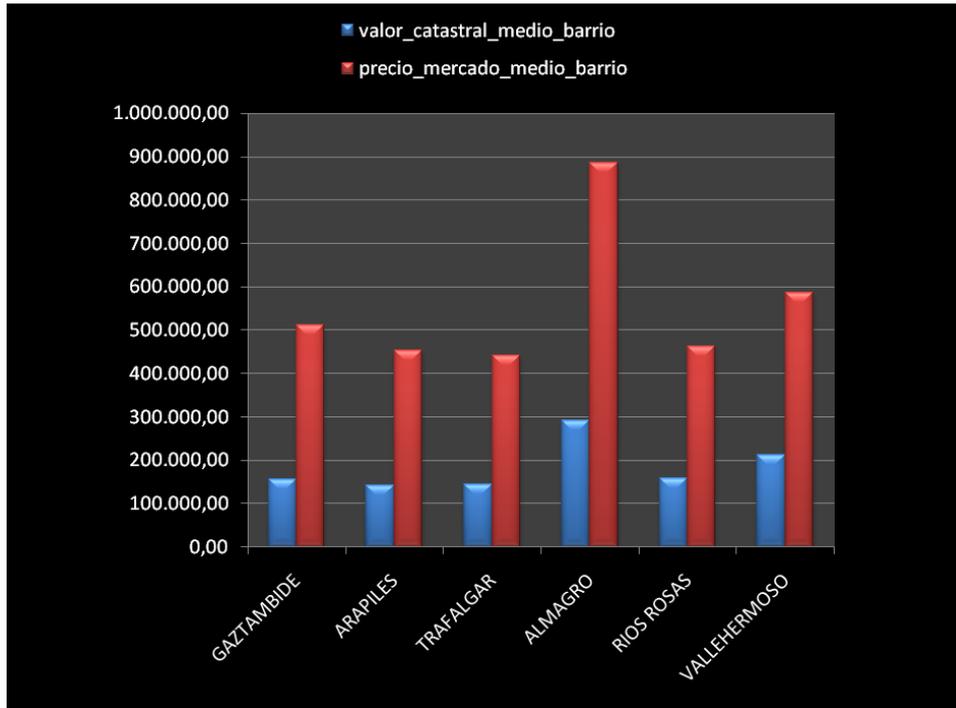
61 Precio de mercado y valor catastral medio de los barrios del distrito Salamanca; fuente: elaboración propia.



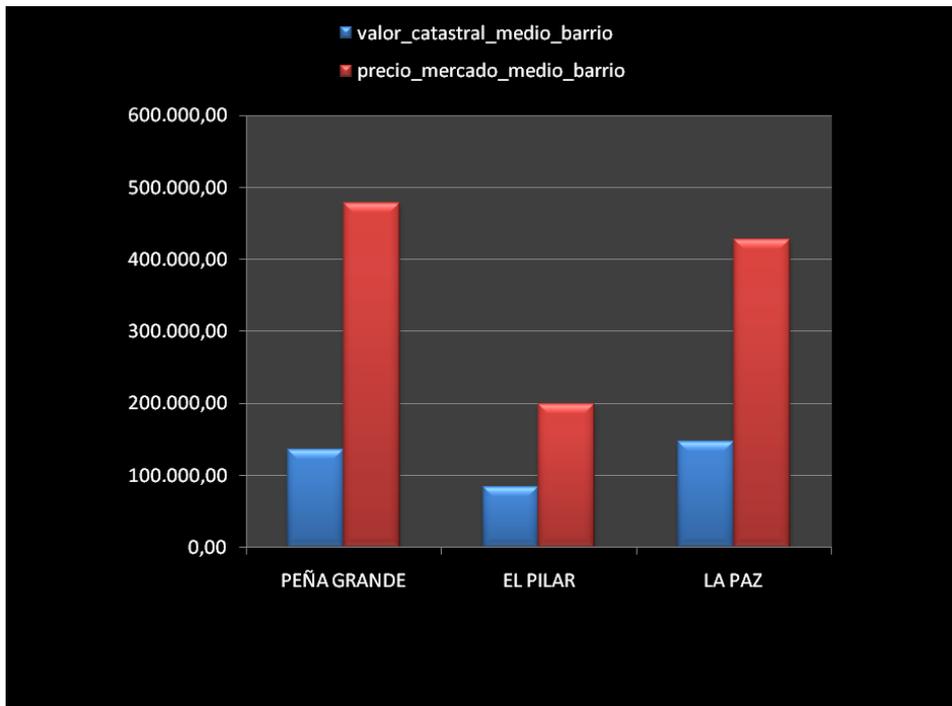
62 Precio de mercado y valor catastral medio de los barrios del distrito Chamartín; fuente: elaboración propia.



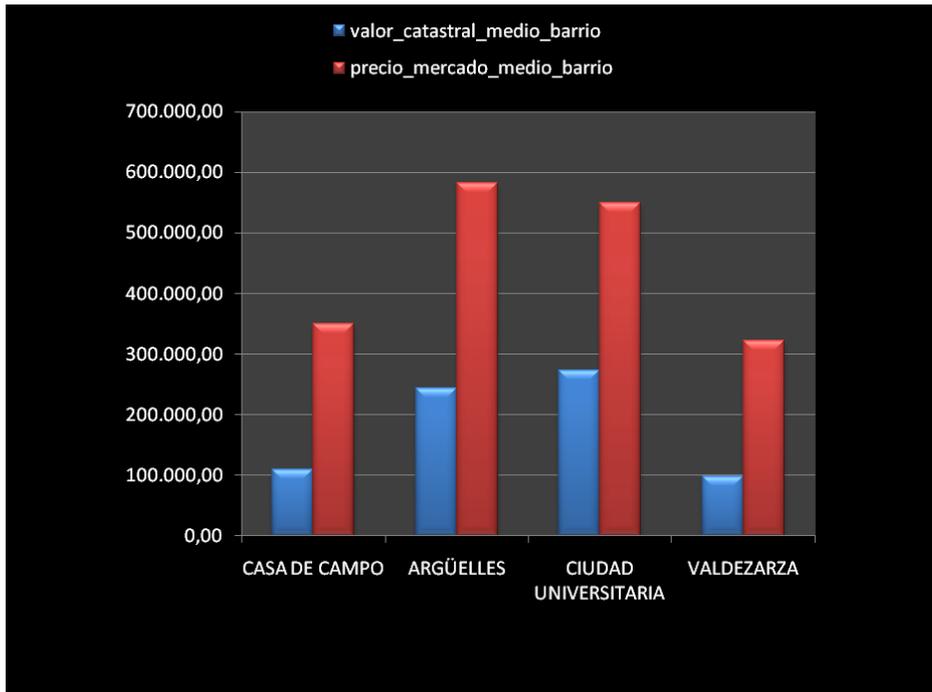
63 Precio de mercado y valor catastral medio de los barrios del distrito Tetuán; fuente: elaboración propia.



64 Precio de mercado y valor catastral medio de los barrios del distrito Chamberí; fuente: elaboración propia.



65 Precio de mercado y valor catastral medio de los barrios del distrito Fuencarral - El Pardo; fuente: elaboración propia.

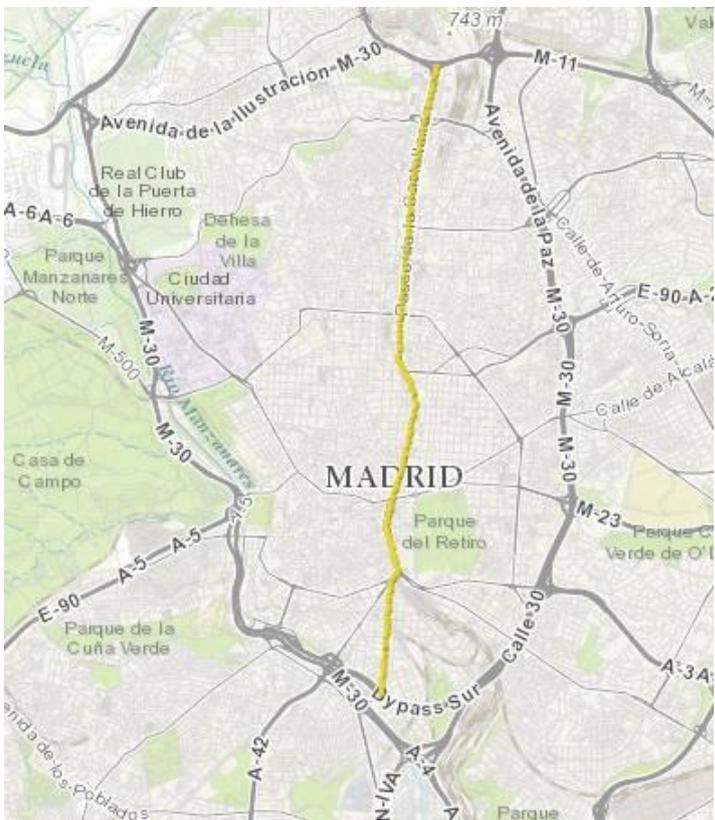


66 Precio de mercado y valor catastral medio de los barrios del distrito Moncloa - Aravaca; fuente: elaboración propia.

## Anexo 2: elementos georreferenciados con el SIG ArcGis online



67 Georreferenciación de la Puerta del Sol; fuente: elaboración propia con ARCGIS online.



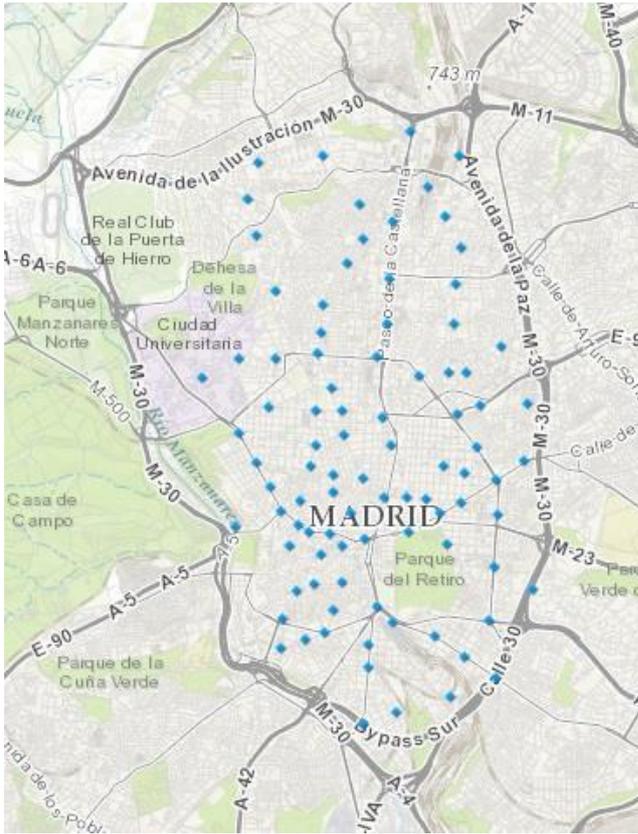
68 Georreferenciación del Paseo de la Castellana; fuente: elaboración propia con ARCGIS online.



69 Georreferenciación de la M30; fuente: elaboración propia con ARCGIS online.



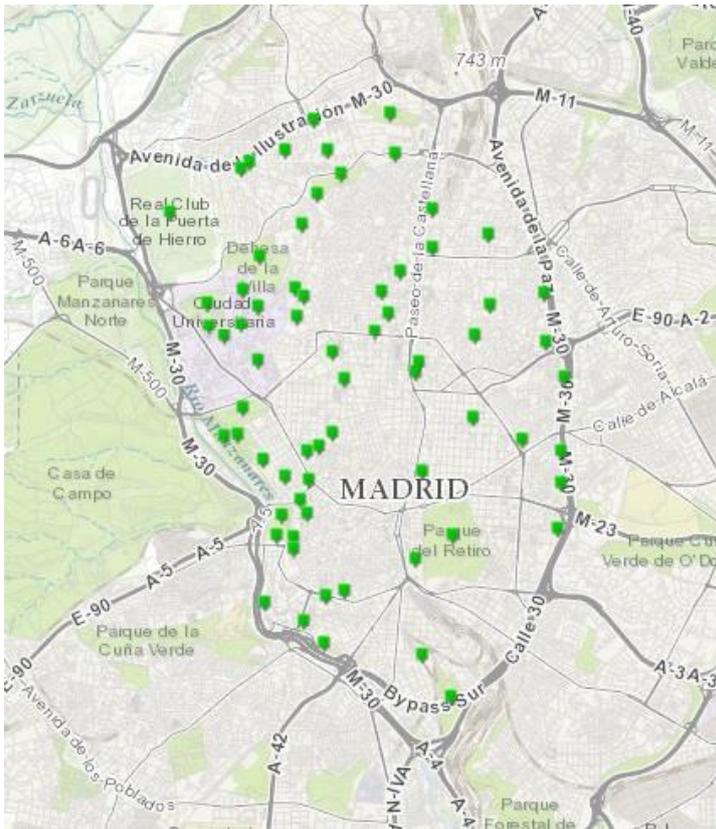
70 Georreferenciación de los nudos de la M30; fuente: elaboración propia con ARCGIS online.



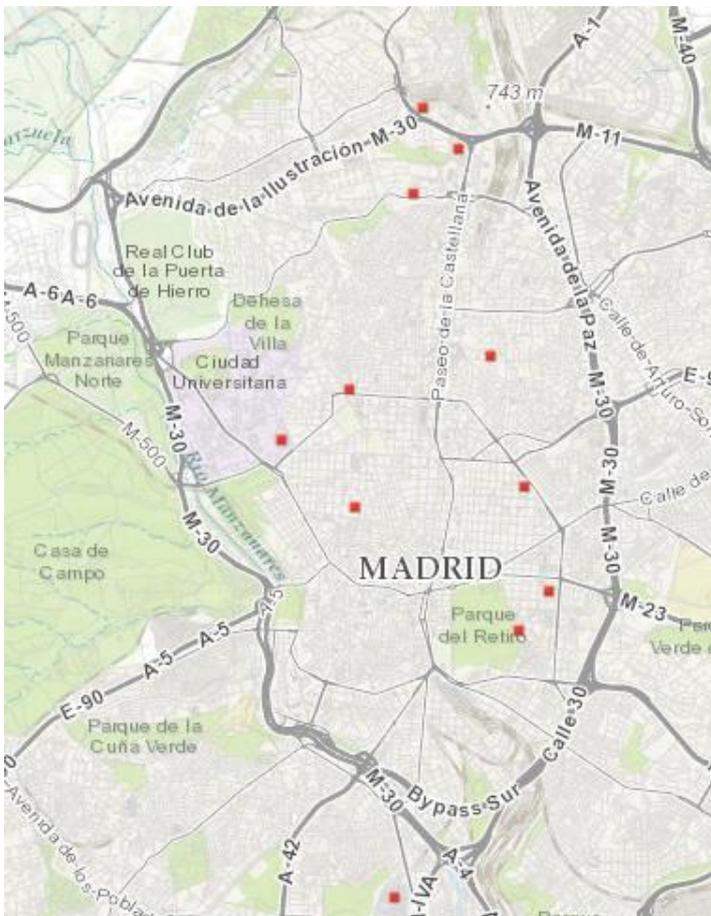
71 Georreferenciación de las estaciones de metro;  
fuente: elaboración propia con ARCGIS online.



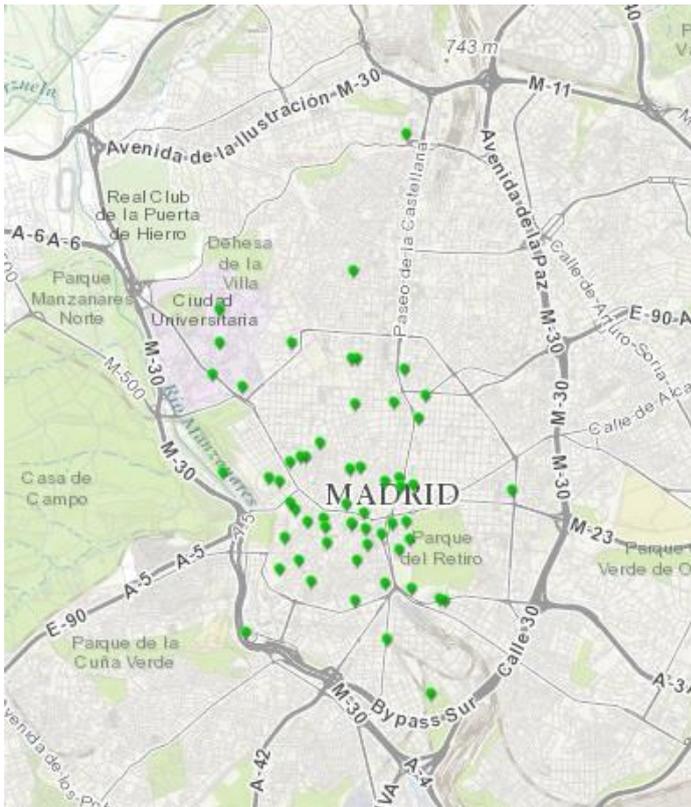
72 Georreferenciación de los colegios públicos;  
fuente: elaboración propia con ARCGIS online.



73 Georreferenciación de las zonas verdes (parques y jardines); fuente: elaboración propia con ARCGIS online



74 Georreferenciación de los hospitales; fuente: elaboración propia con ARCGIS online.



75 Georreferenciación de los museos; fuente: elaboración propia con ARCGIS online.