

MEMORA:

VIVIENDA EN MADRID

Autor: Maria del Rosario, Ortiz

Fecha: 15/12/2023

# 1. Resumen del proyecto

Este es un trabajo de investigación de datos sobre viviendas de Madrid. El objetivo es analizar como se ve afectado el precio por metros cuadrados (€/m2) de una vivienda en Madrid, según distintas variables o características propias de las viviendas como:

* Zona
* Nº de habitaciones
* Balcón/Terraza
* Baños
* Trastero
* Estacionamiento
* Certificado energético

# 2. Introducción

## Contextulización

El mercado inmobiliario en Madrid varia mucho dependiendo de las zonas, por lo que por las mismas caracteristicas en distintos lugares podes observar inmuebles con muchas variaciones de precio.

## Objetivo de la investigación

El objetivo de esta investigación es entender el mercado inmubiliario y ver cuales son las caracteristicas que determinan los precios de los inmuebles.

# 3. Metología

## Recopilación de datos

La fuente de datos utilizada para esta investigación es un data set de Kaggle sobre las viviendas en madrid.

## Procesamiento y limpieza de datos

El data set a analizar tenía muchas filas con valores faltantes, incluso columnas que no tenían ningún valor. Procedí a eliminar solo las columnas que no tenían valores e imputar las filas faltantes.

1. Primero que nada pude ver que ciertas filas solo tenían un valor “True o False” por lo que asumí el faltante era el opuesto.
2. Rellene los valores nulos en la columna de año de construcción en base a si necesitaban reforma o no. También haciendo el análisis univariante pude ver que había un outlier superior a 8000 por lo que lo impute con la media del año de construcción en donde se ubicaba el inmueble.
3. Impute la filas en la columna piso que no tenían valores en base a la moda según la ubicación.
4. Rellene los valores nulos en la columna Bajo tomando como referencia el piso en el que se encuentran.
5. Volví a calcular la columna de metros cuadrados porque tenía valores faltantes, para esto use el precio de venta dividido el precio de venta por m2.
6. Imputé los valores nulos de la columna ‘exterior’ en base al piso en el que se encuentran.
7. Al analizar la columna ‘tipo\_inmueble’ vi a través del anuncio que los valores faltantes eran Estudios por lo que le impute el tipo de vivienda Piso.
8. Rellené los valores nulos en la columna ‘bannos’ (baños) con la media de los baños por metros cuadrados.
9. Después de imputar todos los datos, agregué la columna de zonas, tomando como referencia el barrio en el que se ubican.
10. A su vez para poder graficar un mapa con las ubicaciones, extraje con una librería externa las coordenadas de los barrios.

# 4. Análisis

Comencé analizando gráficamente la variable que me parecía más relevante, que es el precio de venta por metro cuadrado ya que es una forma de estandarizar los precios sin que influya la cantidad de metros cuadrados del inmueble. Luego de haber visto la distribución de dicha variable y concluyendo que no era normal, comencé con las pruebas gráficas para interrelacionar las variables que me parecían más importantes como el CEE, las habitaciones, zonas a donde se ubican los inmuebles y tipo de viviendas. A su vez de estas relaciones se derivan otras cuantas que no eran las centrales originalmente, como lo son tiene o no trastero, estacionamiento, jardín, entre otras. También analicé con pruebas estadísticas no paramétricas ya que mis variables no tienen una distribución normal, si las comprobaciones gráficas eran correctas. Por último transforme las columnas booleandas a numéricas para poder ver un mapa de correlaciones entre dichas variables y realicé un mapa graficando mediante colores en donde se encontraban los inmuebles más caros, como también representando su tamaño en metros cuadrados.

# 5. Resultados

Análisis de la variable principal Precio de venta por metros cuadrados

A graph of a number of people

Description automatically generated with medium confidence

A graph of different colored bars

Description automatically generatedRelación entre el CEE y el precio de venta por metro cuadrado, la letra A es la mejor calificación por lo que se concluye que se está relacionado el precio con esta variable.

A graph of different colored squares

Description automatically generated

Relación entre el tipo de inmueble y el precio de venta por m2, se puede observar que los precios de venta más elevados están en los áticos que son los últimos pisos de los edificios.

A graph with different colored rectangles

Description automatically generatedRelación entre las zonas y el precio por m2, podemos ver que la zona centro de Madrid es la que tiene los precios más elevados.

A graph of different colored bars

Description automatically generated

Relación entre el precio de venta por m2, las zonas y el tipo de inmueble, se puede observar que siempre los áticos son los de mayor precio aunque se cambie de zona.

A graph of a number of dots

Description automatically generated with medium confidence

Relación entre el precio de venta por m2, la cantidad de habitaciones y si tiene estacionamiento.

A graph of different colored dots

Description automatically generated

Relación entre el precio de venta por m2, la cantidad de habitaciones y si tiene jardin.

A graph of a number of dots

Description automatically generated with medium confidence

Relación entre el precio de venta por m2, la cantidad de habitaciones y si tiene trastero.

**Mapa de correlaciones**

A colorful squares with different colors

Description automatically generated

A map with blue dots and red lines

Description automatically generated**Mapa** **de Madrid y los precios por m2**

# **5.** Conclusiones

Podemos concluir que las variables que más afectan a la hora de determinar el precio de un inmueble, son el número de habitaciones, el CEE, el tipo de vivienda y prioritariamente la zona en la que esta ubicada.