

EDA Alojamiento Turístico Madrid

Autores: Patricia García Rabasco, Javier López Cervino, Pilar Gutiérrez Giménez

Bootcamp Data Analytics — Memoria del Proyecto

1. Introducción

En los últimos años, el alquiler turístico se ha convertido en un fenómeno global que ha modificado significativamente el mercado inmobiliario y la oferta de alojamiento en ciudades de todo el mundo. Madrid, como una de las capitales europeas con mayor afluencia turística, no es una excepción. Plataformas como Airbnb permiten a propietarios y gestores ofrecer alojamiento de manera flexible, lo que ha impulsado tanto oportunidades económicas como debates sobre su impacto en la vivienda, el turismo y la convivencia vecinal.

Este proyecto tiene como objetivo realizar un análisis exploratorio de datos (EDA) sobre los alojamientos turísticos en Madrid utilizando un conjunto de datos real obtenido de Inside Airbnb. Mediante este análisis buscamos entender mejor cómo se distribuyen los alojamientos dentro de la ciudad, qué características comparten, cómo varían sus precios y qué factores pueden influir en estas variaciones.

La memoria que se presenta a continuación recoge el proceso completo: desde el planteamiento inicial y la hipótesis, hasta los resultados obtenidos y las conclusiones del estudio. Asimismo, se plantea trabajo futuro para continuar desarrollando el análisis más allá de la fase exploratoria.

2. Objetivos del análisis

El análisis se plantea con una serie de objetivos que guían la exploración de los datos. Estos objetivos permiten estructurar el trabajo y centrar los esfuerzos en preguntas relevantes para comprender el funcionamiento del mercado de alojamientos turísticos en Madrid.

- Analizar la distribución geográfica de los alojamientos para identificar qué barrios concentran mayor oferta turística.
- Estudiar la variación de precios por zonas y por tipo de alojamiento.
- Investigar qué características del alojamiento (tipo, disponibilidad, número de reseñas, etc.) podrían influir en el precio.
- Comparar el comportamiento de anfitriones particulares con el de anfitriones que gestionan múltiples propiedades (multi-hosts).
- Revisar patrones de disponibilidad y posibles indicadores de demanda.

3. Hipótesis del proyecto

La hipótesis principal del proyecto propone que:

“Las diferencias de precio entre los alojamientos turísticos de Madrid pueden explicarse principalmente por factores como el barrio, el tipo de alojamiento y las características del anuncio.”

Esta hipótesis surge de la observación de tendencias comunes en mercados similares, donde la ubicación suele ser un elemento determinante en el valor de un inmueble y donde las características del alojamiento (como si es un piso completo o una habitación privada) añaden diferenciación significativa al precio final.

A lo largo del análisis exploratorio se examinan distintas variables que permiten confirmar o matizar esta hipótesis.

4. Descripción del dataset

El dataset utilizado procede de Inside Airbnb, una plataforma que proporciona datos públicos acerca de los alojamientos disponibles en diversas ciudades del mundo. En este caso se ha trabajado con el archivo listings.csv correspondiente a la ciudad de Madrid.

- Registros totales: 25.094
- Columnas: 18

Entre las variables disponibles se encuentran: identificador del alojamiento, nombre del anuncio, barrio, latitud y longitud, tipo de alojamiento, precio por noche, estancia mínima, número de reseñas, disponibilidad anual y otros atributos relevantes.

El dataset permite estudiar múltiples dimensiones del mercado turístico: desde la distribución espacial hasta el comportamiento de los anfitriones.

5. Metodología

El proceso metodológico seguido para realizar el EDA se basa en una combinación de técnicas de análisis descriptivo, visualización de datos y limpieza y estructuración del dataset. Las etapas principales fueron:

1. Limpieza del dataset: conversión de precios a valores numéricos, eliminación de duplicados, tratamiento de valores nulos y casos extremos.
2. Análisis descriptivo: revisión estadística básica de variables numéricas y categóricas, observación de distribuciones y detección de patrones generales.
3. Visualización de datos: generación de gráficos y mapas para interpretar la distribución geográfica, la variación de precios y otros aspectos clave.
4. Comparación entre grupos: estudio de diferencias entre tipos de alojamiento, barrios y anfitriones.
5. Interpretación de resultados: análisis de tendencias relevantes en el contexto del mercado de alojamientos turísticos.

6. Resultados del análisis

Los principales resultados del análisis permiten obtener una visión clara del mercado de alojamientos turísticos en Madrid. Entre las conclusiones más destacadas se encuentran las siguientes:

- **Distribución geográfica:** la mayor concentración de alojamientos se encuentra en los barrios del centro de la ciudad, seguidos por Salamanca, Chamberí y Retiro. Las zonas periféricas tienen menor presencia de anuncios.
- **Precios:** los alojamientos situados en zonas céntricas presentan precios significativamente superiores. También se observa una diferencia notable entre pisos completos y habitaciones privadas, siendo los primeros más caros.
- **Disponibilidad:** la disponibilidad anual es menor en zonas más turísticas, lo que sugiere una mayor demanda.
- **Multi-hosts:** un porcentaje significativo de los anuncios pertenece a anfitriones que gestionan varias propiedades, lo que indica cierta profesionalización dentro del mercado.
- **Características del alojamiento:** variables como el número de reseñas o la estancia mínima muestran correlaciones moderadas con el precio, aunque menos influyentes que la ubicación y el tipo de alojamiento.

7. Conclusiones

El análisis realizado confirma la hipótesis principal del proyecto: los precios de los alojamientos turísticos en Madrid están fuertemente condicionados por la ubicación y el tipo de alojamiento. Las zonas céntricas, debido a su atractivo turístico, presentan precios más altos y menor disponibilidad anual, lo que puede interpretarse como un indicador de alta demanda.

Asimismo, se observan diferencias sustanciales entre anfitriones particulares y multi-hosts, siendo estos últimos responsables de una parte importante de la oferta total, especialmente en las zonas de mayor demanda.

En conjunto, el estudio permite comprender mejor la dinámica del mercado de alojamiento turístico en Madrid, aportando información valiosa tanto para gestores profesionales como para análisis urbanísticos y sociales.

8. Trabajo futuro

Este proyecto abre múltiples posibilidades para análisis más avanzados. Algunas propuestas de trabajo futuro incluyen:

- Desarrollar modelos predictivos que permitan estimar el precio del alojamiento a partir de sus características.
- Realizar análisis temporal utilizando datasets de años anteriores para estudiar la evolución del mercado.
- Incorporar análisis de sentimiento a partir de reseñas de usuarios, en caso de disponer de estos datos.
- Ampliar el estudio comparando Madrid con otras ciudades españolas o europeas para contextualizar los resultados.
- Estudiar la relación entre la presencia de Airbnb y variables socioeconómicas de los barrios.