



**Universidad  
Europea**

**UNIVERSIDAD EUROPEA DE MADRID**

**ESCUELA DE ARQUITECTURA, INGENIERÍA Y DISEÑO**

**GRADO EN MÁSTER DATA SCIENCE**

**ANTEPROYECTO**

**PREDICCIÓN INTELIGENTE DE PRECIOS  
EN AIRBNB MADRID**

**KATHERINE LÓPEZ RAMÍREZ**

**CURSO 2024 - 2025**



**TÍTULO:** PREDICCIÓN INTELIGENTE DE PRECIOS EN AIRBNB MADRID

**AUTOR:** KATHERINE LÓPEZ RAMÍREZ

**TITULACIÓN:** GRADO EN MÁSTER DATA SCIENCE

**DIRECTOR DEL PROYECTO:** ÁLVARO SÁNCHEZ PÉREZ

**FECHA:** ABRIL DE 2025

# Índice

1. JUSTIFICACIÓN .....	5
2. OBJETIVOS DEL PROYECTO.....	5
1.1    Objetivo general .....	5
1.2    Objetivos específicos .....	5
1.2.1    Dentro del alcance .....	6
1.2.2    Fuera del alcance .....	6
3. PLAN DE PROYECTO.....	6
REFERENCIAS BIBLIOGRÁFICAS.....	9

# 1. JUSTIFICACIÓN

El mercado de alojamientos turísticos ha experimentado un gran auge en la última década, siendo plataformas como Airbnb protagonistas del cambio. Madrid, como capital turística, representa un entorno diverso en cuanto a oferta y demanda de alojamientos a corto plazo. En este contexto, conocer y optimizar los precios de publicación se vuelve clave para anfitriones que buscan maximizar sus ingresos en un entorno cada vez más competitivo.

Este proyecto se plantea como una herramienta para entender los factores que determinan el precio de las propiedades listadas en Airbnb, haciendo uso de técnicas de análisis de datos y modelado predictivo aplicadas a datasets reales.

El objetivo es predecir los precios de alojamientos en función de zonas geográficas y estacionalidad turística, con especial énfasis en:

- La estacionalidad turística (días de mayor demanda frente baja ocupación)
- La predicción de precios a partir de variables clave
- La posibilidad de incorporar soluciones en tiempo real mediante scraping o conexión a APIs, si resulta viable

Inspirado en modelos utilizados por plataformas como PriceLabs y AirDNA, el proyecto permitirá ofrecer una solución sencilla, visual y práctica para que anfitriones puedan entender y ajustar sus precios según el contexto del mercado. Este trabajo permitirá aplicar conocimientos en análisis, modelado, visualización y comunicación de resultados, conectando teoría con práctica real.

## 2. OBJETIVOS DEL PROYECTO

### 1.1 Objetivo general

Analizar y predecir los precios de alojamientos en Airbnb Madrid en función de la zona geográfica y la estacionalidad turística, aplicando técnicas de machine learning.

### 1.2 Objetivos específicos

- Obtener y preparar datos de Airbnb en Madrid desde Inside Airbnb.

- Identificar fechas de alta y baja demanda turística y su relación con los precios.
- Analizar diferencias de precios por zona geográfica.
- Construir modelos de predicción de precios utilizando técnicas de regresión y aprendizaje automático (regresión lineal, random forest, XGBoost...).
- Evaluar modelos y validar métricas de rendimiento.
- (Opcional) Conectar con datos dinámicos mediante scraping o API si es viable.
- Desarrollar una visualización interactiva del modelo predictivo.
- Comparar los resultados con plataformas comerciales como PriceLabs y AirDNA.

### 1.2.1 Dentro del alcance

- Análisis exploratorio y descriptivo completo.
- Visualización geográfica y por categorías.
- Predicción de precios con Machine Learning.
- Comparación con soluciones comerciales (PriceLabs, AirDNA).

### 1.2.2 Fuera del alcance

- Desarrollo de una aplicación completa en producción.
- Predicciones en tiempo real con datos 100% actualizados (si scraping no es viable en el tiempo disponible).
- Segmentación por tipo de alojamiento (apartamento, estudio, etc.)
- Segmentación por capacidad o número de personas.

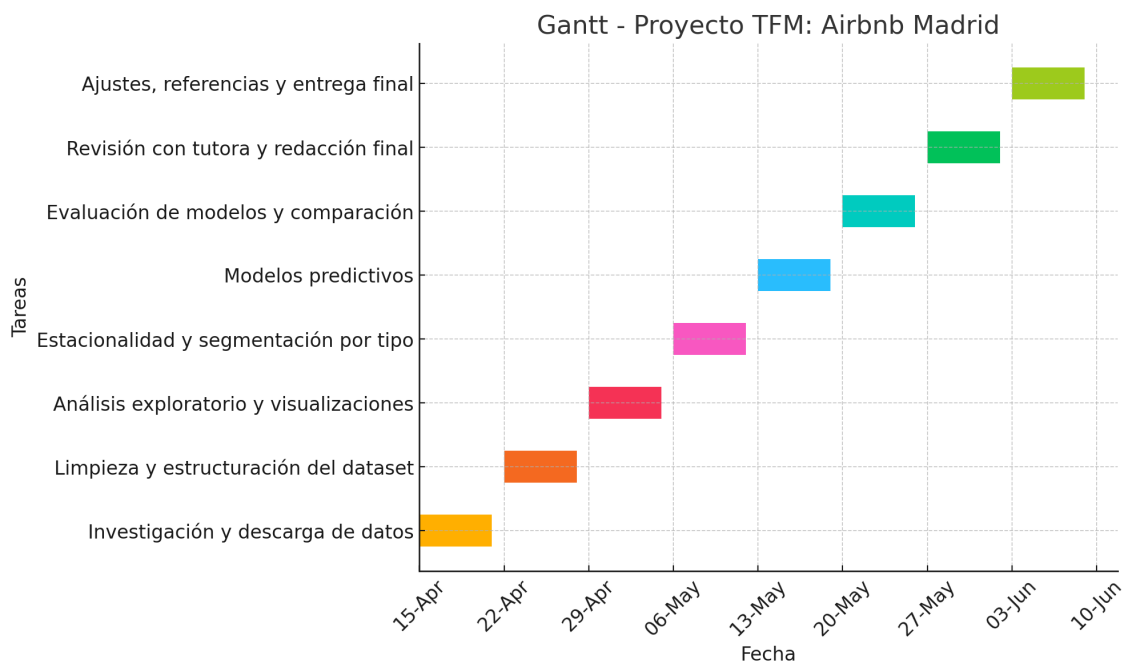
## 3. PLAN DE PROYECTO

### *Fases*

1. Documentación y descarga de datos
  - Inside Airbnb, AirDNA, PriceLabs
  - Posible scraping de webs si es viable (opcional)
2. Preparación del dataset
  - Limpieza y preparación de variables clave (zona, precio, disponibilidad, fecha, reviews)
  - Geocodificación y categorización por zonas

3. Análisis exploratorio y visualización
  - Distribución de precios por distrito
  - Mapa interactivo y gráficos
4. Estacionalidad y series temporales
  - Agregación por fecha de calendario
  - Cruce con eventos turísticos
5. Modelado predictivo de precios
  - Pruebas con regresión, (random forest, XGBoost) o modelos de clustering
  - Validación de modelos y visualización de resultados
  - (Opcional) Exploración de scraping/API
  - Creación de visualización interactiva del modelo
6. Informe y conclusiones
  - Comparación con plataformas comerciales
  - Recomendaciones prácticas
  - Documentación técnica
7. Revisiones con tutora y entrega final

Fase	Tarea principal	Duración estimada
F1	Investigación y descarga de datos (Inside Airbnb, PriceLabs, AirDNA)	Semana 1
F2	Limpieza y estructuración del dataset	Semana 2
F3	Análisis exploratorio de datos + visualizaciones	Semana 3
F4	Estacionalidad turística y análisis por zonas geográficas	Semana 4
F5	Modelos predictivos (regresión, árboles...)	Semana 5
F6	Evaluación de modelos y comparación con plataformas	Semana 6
F7	Revisión con tutora y escritura del informe final	Semana 7
F8	Ajustes, referencias y entrega final	Semana 8





## REFERENCIAS BIBLIOGRÁFICAS

- Inside Airbnb. <https://insideairbnb.com/>
- PriceLabs. <https://hello.pricelabs.co/>
- AirDNA. <https://www.airdna.co/>
- Airbnb Help Center, Pricing & Hosting Best Practices.