



PROYECTO

1. Definir Data Set

El dataset escogido es [éste](#), scrapeado de Airbnb. Vuestra primera misión será familiarizaros con lo que ahí se recoge.

2. Arquitectura y validación de los datos

- a. Muestreo y exploración inicial de los datos
- b. Definir e implementar el Datawarehouse
- c. (Opcional) Ingesta de datos (ETL) y validación de que se ha cargado correctamente

3. Análisis Exploratorio

Hacer un estudio estadístico con R o Python, según preferencia personal, y averiguar cuales son las métricas adecuadas para el dataset. No olvidemos:

- a. Revisión de la calidad de los datos
- b. Detección outliers (rango de variables), imputación valores nulos.
- c. Boxplots, histogramas, etc.
- d. Normalización de los valores de las tablas (quitar tildes, “dobles espacios”, etc.)

4. Visualización de las métricas

A partir de los datos de Airbnb, obtén los KPIs que puedan ser de relevancia y contesta a través de un dashboard a una pregunta relevante que hagas sobre los datos.

- a. a. Se valorará el diseño final del dashboard.
- b. b. El uso de buenas prácticas.
- c. c. El cálculo de KPIs adecuados y el uso de campos calculados avanzados.
- d. d. El uso de vistas interactivas.

Nota: En este ejercicio no habrá un dashboard modelo sino que se basará en valorar vuestras capacidades con el uso de la herramienta de Tableau y que podáis demostrar todo lo aprendido durante este bloque.

5. Pre-procesamiento y Modelado

La tarea asignada es hacer un algoritmo de regresión lineal que prediga el precio de un inmueble en función de las características que elijáis.

6. Informe

En esta etapa se debe de simular la presentación de resultados en un entorno real de empresa.

- Suposiciones iniciales
 - Cuales han demostrado ser válidas y cuáles no. ¿Por qué?
- Métricas seleccionadas: ¿han sido las correctas o no? ¿por qué?

KeepCoding© All rights reserved.

www.keepcoding.io



- Teniendo en cuenta lo aprendido ¿Qué cosas se harían igual y cuales se harían de otra forma? ¿Por qué?
- Conclusiones y “lessons learned”

-