

Trabajo:

Para el dataset "SmartphoneMessy", es necesario realizar un Análisis Exploratorio de Datos (EDA) que debe incluirse en un notebook llamado "Nombre_Apellido.ipynb".

Este análisis deberá contener los siguientes elementos:

- Importación de las librerías necesarias.
- Lectura y carga del conjunto de datos.
- Exploración inicial de los datos.
- Análisis de la cardinalidad de las variables.
- Tratamiento de valores faltantes y anomalías.
- Se puede optar por separar los datos en dos dataframes, especialmente para segmentar según las NaNs en la columna 'rating'.
- Por ejemplo, convertir los precios a euros, separar la información en la columna 'sim', procesar la columna 'ram', etc.
- Exploración de cada columna individualmente.
- Estudio de las relaciones entre las variables más relevantes y la variable objetivo ('Rating').
- Utilizar el modelo OLS de Statsmodels (<https://www.statsmodels.org/stable/examples/notebooks/generated/ols.html>)
- LinearRegression de SciKit-Learn (https://scikit-learn.org/stable/modules/generated/sklearn.linear_model.LinearRegression.html) para imputar o predecir los valores faltantes.