BREVE RECORDATORIO:

Pasos del EDA

1. Carga y limpieza inicial del dataset:

- Cargar los datos proporcionados (descargar desde el enlace provisto).

- Identificar y documentar la estructura del dataset.

- Verificar la presencia de valores faltantes y determinar su tipo (MCAR, MAR, MNAR).

2. Análisis exploratorio de datos (EDA):

- Crear gráficos y tablas que describan el conjunto de datos y que ayuden a visualizar problemas de calidad, incluyendo:

- Valores faltantes

- Desbalanceo de clases

- Posibles sesgos en los datos

3. Imputación de datos faltantes:

- Aplicar diferentes técnicas de imputación (eliminar filas, imputación por la media, regresión, etc.).

- Comparar los resultados obtenidos con diferentes métodos y justificar la elección del método final.

4. Manejo del desbalanceo:

- Implementar técnicas de muestreo (sobre/submuestreo o SMOTE) para lidiar con clases desbalanceadas en el dataset.

- Evaluar el impacto del desbalanceo en los modelos predictivos.