Tarea 1: Preprocesamiento de Datos

- **Propósito:** Introducir a los estudiantes al manejo y limpieza de datos. Desarrollar habilidades para identificar datos faltantes, valores atípicos y escalado de variables, fundamentales para el modelado.
- Herramientas a utilizar:
 - o Google Colab: para codificación y documentación del análisis.
 - Lenguaje R: con librerías como ggplot2, VIM, DMwR2, psych, scales, etc.
 - o **Github:** Guardar base de datos en su repositorio
- Descripción de la actividad:
 - Se entregará un conjunto de datos en formato .csv con inconsistencias comunes (valores faltantes, duplicados, diferentes formatos de fechas, etc.).
 - o Los estudiantes deben:
 - 1. Realizar un análisis exploratorio de datos, con las funciones, dim(), str(), describe(), hist(), bartplot(), pie(). **1punto**
 - 2. Identificar y manejar valores faltantes, con gráficos utilizados y matrices. **1punto**
 - 3. Realizar imputación de valores utilizando métodos como la media, moda o regresión, y métodos robustos. **1punto**
 - 4. Detectar y manejar valores atípicos univariados y multivariados de las variables numéricas. **1punto**
 - 5. Normalizar/escalar las variables numéricas presentar en un solo gráfico los boxplot de las variables numérica, y todas las transformaciones trabajadas en clases. **1punto**
 - 6. Documentar cada paso en Google Colab. 2puntos
 - 7. Un video con cámara encendida de máximo 5 minutos. de la explicación del trabajo, propósito, código en colab, base de datos en github, y los resultados obtenidos. **3puntos**
 - o **Resultado esperado:** Un conjunto de datos limpio, listo para el análisis.

Nota: Subir un archivo PDF con portada y sólo agregar los enlaces de Google Colab y del video.

Base de datos: Patrones de carga de vehículos eléctricos

Conjunto de datos con 1,320 sesiones de carga, incluyendo consumo de energía, duración y detalles del vehículo. Permite analizar el comportamiento del usuario y modelar patrones predictivos. https://www.kaggle.com/datasets/valakhorasani/electric-vehicle-charging-patterns

Minería de Datos: Emanuel G. Muñoz