

# Description Report (Patadas)

- **Estructura del dataset**
  - El dataset parece seguir una estructura de tipo **Panel Data**. Donde se miden muchas variables de todas las vacas en diferentes horas.
    - El dataset contiene 37 features.
    - El número de instancias es de 43.
    - El periodo de los registros compete a las horas del día 18/07/2025.
- **Calidad de los datos**
  - El dataset tiene diversos tipos de variables:
    - Datos de tipo Int y Float64
    - Objects como variables categóricas
    - Las únicas variables de tipo categóricas corresponden a los objetos que son los siguientes:
      - Hora Inicio Ordeño (fecha en formato DD-MM-YYYY HH:MM)
      - Media Interv. Ordeño (Hora en formato HH:MM)
    - Features donde todos sus registros parecen ser nulos
- **Recomendaciones**
  - Imputar datos (dependiendo del contexto) con 0's, promedio o eliminar completamente los registros.
  - Modificar el tipo de datos numéricos a Float64 para evitar conflictos entre Int y Float.
  - Se necesita verificar a que hace referencia cada lactancia, ya que no tienen que ver con producción (para esto último se tiene cada dataset individual de cada vaca para corroborar este dato).
  - Usar DEL (Días en leche, cuántos días lleva produciendo leche) como variable pivote para encontrar correlaciones entre las demás variables; ya que esta variable parece que tendrá una gran relevancia para nuestro proyecto.