

# Description Report (Ficha Vaca)

- **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 3512 features
    - El número de instancias es de 7
- **Calidad de los datos**
  - El dataset presenta principalmente variables de tipo texto (object), ya que la mayoría de los campos corresponden a descripciones, fechas o nombres de usuarios.
  - No existen variables numéricas útiles, salvo una columna (**Evento**) de tipo float64, aunque todos sus registros son nulos.
  - Las variables categóricas corresponden a los campos con valores repetitivos o de clasificación:
    - Tipo de evento → Categorías como “Diagnósticos/Tratamiento”, “Cambio tabla alimentación”, “Invitación Visita Veterinaria”, entre otros.
    - Usuario → Roles o responsables que ejecutan los eventos (por ejemplo, veterinarios o administradores).
  - Fecha del evento está en formato texto (DD/MM/YYYY), por lo que requiere conversión a tipo **datetime** para análisis temporales.
  - La columna “Comentario” contiene una gran cantidad de valores nulos (>80%), y “Evento” está completamente vacío.
  - En general, existen features con registros nulos o información redundante, lo que sugiere la necesidad de depuración previa antes de realizar análisis descriptivos o inferenciales.
- **Recomendaciones**
  - Eliminar columnas irrelevantes o completamente nulas, como **Evento**, ya que no aportan información al análisis.
  - Depurar y estandarizar las fechas en **Fecha del evento** al formato **datetime (YYYY-MM-DD)** para facilitar análisis temporales o filtrado por períodos.
  - Clasificar los tipos de eventos en categorías más generales (por ejemplo: *salud, alimentación, manejo, reproducción*), lo que permitirá detectar patrones o concentraciones de actividades.
  - Revisar los valores nulos en la columna **Comentario**, si los comentarios son informativos, podrían utilizarse para análisis cualitativos (palabras clave, frecuencia de términos). Si no, pueden eliminarse.

- Verificar la consistencia de usuarios (por ejemplo, veterinarios vs administradores) para evitar duplicados por diferencias de escritura o formato.
- En caso de integrarse con otros datasets (por ejemplo, producción o historial de vacas), utilizar la fecha del evento como variable pivote para cruzar información y analizar impacto de los eventos en la productividad o bienestar.
- Realizar una codificación categórica (Label o One-Hot Encoding) sobre **Tipo de evento** y **Usuario** si se desea aplicar modelos predictivos o análisis multivariado posteriormente.