

Transformaciones (Tercera iteración)

Dataset de Fichas de Vaca (HATO)

La hora registrada en el dataset original se encuentra en formato string (u object), por lo que una transformación esencial consiste en convertir en tipo datetime y estandarizar su nombre a “Fecha”, en lugar de mantenerlo como “Fecha del Evento”.

En caso de encontrarse fechas inválidas, sin datos registrados o que no contengan valores numéricos, dichas entradas se clasifican y almacenan como fichas inválidas para su análisis posterior.

Dataset de Registros de ordeño (HATO)

Los campos de fecha y hora requieren transformación, ya que originalmente se encuentran integrados dentro de la columna Hora de inicio, la cual contiene formatos inconsistentes. En primer lugar, se normaliza el formato temporal, reemplazando expresiones como “a. m.” y “p. m.” por los formatos estándar AM y PM. Posteriormente, la columna Hora de inicio se convierte a tipo datetime, lo que permite extraer de forma independiente la fecha y la hora exacta de inicio del ordeño.

Además, es necesario transformar las columnas que registran valores en formato minutos-segundos o horas-minutos a una métrica uniforme expresada únicamente en minutos. Esta normalización garantiza que las duraciones del ordeño y los intervalos entre ordeños sean comparables y analizables de manera consistente dentro del dataset.