

Data Exploration & Preparation Archivos CSV

Información General

De acuerdo al entendimiento de los datos que se observa en la documentación de [Rationale for inclusion/exclusion](#) podemos deducir que trabajaremos con el datos de registros de ordeño, ya que tiene información relevante para cumplir con los [Business objectives](#).

ID Vaca(1204-8794)

https://colab.research.google.com/drive/1GSU7paZfLqE5-2Y_LX-T_lku8YlKIKKG?usp=sharing

Reporte

https://colab.research.google.com/drive/1AkoDETONsIBZZgHY2hN62S6JvACSW5_Z?usp=sharing

Inventario

https://colab.research.google.com/drive/1T_WFIti2eqRdEr_kaHqCquC9gDIchke7?usp=sharing

Patadas

https://colab.research.google.com/drive/1Bp0E02UkeJ_Fn8-pj04_OEsfi-oQHItN?usp=sharing

Datos de ordeña - Todas las vacas

https://colab.research.google.com/drive/1X_sfQIY68aH8Z1Cru19x5Cajw21sVAXx?usp=sharing

Pipeline para limpiar

MilkingDataPipeline es un pipeline de procesamiento de datos para análisis de producción lechera que implementa las fases de Comprensión de Datos y Preparación de Datos de CRISP-DM.

Objetivos del Pipeline:

- Preparar datos históricos de ordeño para análisis predictivo
- Normalizar variables con diferentes escalas y distribuciones
- Crear features relevantes para modelado de series temporales

[Pipeline](#)