

Endometriosis Dataset — Video Summary (3 Phases)

Author: Vishal

What the videos cover: A focused walkthrough of how I audited, cleaned, engineered, and validated the dataset to produce a modeling-ready table, plus the key clinical takeaways that guide next steps.

Phase 1 — Initial Dataset Audit & Cleanup

- What it is: A zero-change baseline review and hygiene pass to make the data trustworthy.
- Output: Clean base tables and a clear split into numeric vs categorical tracks.

Phase 2 — Numerical Analysis & Feature Engineering

- What it is: Focus on quantitative variables to build timelines and structure QoL signals.
- Output: Modeling-ready numeric table with engineered features and validated imputations.

Phase 3 — Categorical Deep-Dive

- What it is: Structured treatment of categorical fields to extract reliable signal without leakage.
- Output: Encoding-ready categorical table with curated levels and association map.

The video can be accessible by this link:

 [vishal_Video_expaintion.mp4](#)