

A summary report for Task 1

Data Tagging

For Task 1, I first read each complaint, cause, and correction carefully to understand what happened. I then matched the information to the closest terms in the provided taxonomy. To keep the tagging consistent, I used **data validation** with the taxonomy lists, which helped me select only approved labels and avoid accidental variations.

For **Root Cause**, I chose the term that best represented why the issue occurred based on the cause text. For **Symptom Condition**, I focused on what the operator or technician observed (leaks, missing parts, broken items, fault codes, etc.). For **Symptom Component**, I identified the part mentioned in the complaint, and when it wasn't available in the taxonomy, I used "Not Mentioned" as required. For **Fix Condition** and **Fix Component**, I mapped the actions in the correction such as replacing, repairing, installing, or tightening—to their closest taxonomy terms.

Insights:

Across the dataset, many issues point to assembly and quality control gaps—loose fittings, missing components, poor threading, and leaking seals. There is also a recurring pattern of sensor-related faults, often caused by contamination or alignment issues. These trends suggest the need for improved assembly checks, better sealing practices, and stronger pre-delivery inspections to reduce repeat failures.