Task 1 – Data Tagging Summary Report

This report summarizes the approach used to tag complaint records into structured categories: Root Cause, Symptom Condition, Symptom Component, Fix Condition, and Fix Component. The objective was to standardize free-text descriptions for effective analysis.

# Tagging Methodology

We analyzed three fields — Complaint, Cause, and Correction — using a rule-based script. Specific keywords and contextual phrases were matched against a predefined taxonomy. For example, terms like “not tightened” or “leaking” were mapped to standardized root causes. When clear matches were missing, we assigned “Not Mentioned” to avoid incorrect assumptions.  
  
Symptom Conditions such as “won’t stay open” or “crushed” were identified from customer complaints. This step required interpretation, as language varied across entries. Similarly, components mentioned in complaints were matched or inferred using known part names or contextual clues.  
  
Corrections were tagged based on action words like “replaced,” “tightened,” or “repaired.” These indicated the nature of the fix and were also used to identify the component involved, if specified. In cases where corrections were vague or missing, tagging was marked accordingly.

# Key Findings

- Assembly-related issues were more common than part failures, suggesting a need for improved installation processes.  
- Components like hydraulic systems and door mechanisms appeared frequently, pointing to potential design or supplier concerns.  
- Most fixes involved replacements, which may reflect a preference for quick resolution or gaps in repair protocols.  
- Several components showed repeat failure patterns, indicating potential for predictive maintenance strategies.

# Recommendations

Future improvements could include NLP-based enhancements such as contextual phrase matching or similarity-based tagging, which go beyond the current rule-based approach.