Data Tagging Summary – Task 1

1. Approach

The tagging process was carried out using the original dataset (DA – Task 1) and structured into tagged_data_main.csv with the help of a predefined taxonomy. The steps were:

• Root Cause Tagging: Extracted the underlying reasons from the Cause column, standardizing variations such as "Not tighten at factory" into normalized categories like not tighten or not installed.

• Symptom Tagging: The Complaint field was analyzed for descriptive terms (e.g., loose, leaking, missing, crushed) and mapped to affected parts such as fuel door, compressor line, bulkhead connector. Multiple issues were separated into distinct condition–component sets.

• Fix Tagging: The Correction column provided action verbs like installed, retightened, replaced, which were mapped under Fix Condition, while the corresponding part (e.g., gas strut, O-ring, bolts) was tagged as Fix Component. This captured the repair actions and focus areas.

2. Challenges

Several challenges were encountered during tagging:
• Many complaints were vague, limiting precise classification.
• Multi-symptom and multi-fix entries required careful separation to avoid overlap.
• Inconsistent phrasing, shorthand, and spelling variations required preprocessing to maintain accuracy.

3. Key Insights

• Frequent Root Causes: The most common tags were not installed and not tightened, pointing toward assembly and installation gaps.
• Recurring Complaints: Frequent mentions of leaks, missing parts, and crushed lines highlight quality control issues in assembly and materials.
• Fix Patterns: The majority of fixes involved simple installation or retightening, suggesting that many failures could be prevented through improved QA and assembly checks.
• Data Gaps: Around 15–20% of entries lacked sufficient detail for tagging, highlighting the need for standardized and complete complaint reporting.

4. Implications

This tagging exercise provides valuable insights into recurring issues and repair trends. By improving assembly protocols, enforcing technician training, and standardizing complaint documentation, manufacturers can reduce repeat complaints and improve product reliability.