PROJECT REPORT

FOR TASK 1

Summary Report on Tagging Process

Used Python and Pandas to automate tagging for the following fields:
Root Cause, Symptom Condition, Symptom Component, Fix Condition, and Fix Component.

Step 1 – Data Cleaning

- Removed extra spaces and special characters.
- Converted all text to lowercase for uniformity.
- Standardized column names in both Task and Taxonomy sheets.
- This ensured accurate comparisons and minimized mismatches.

Step 2 – Matching Technique

- Implemented fuzzy string matching using the function fuzzywuzzy.process.extractOne.
- This method finds the closest matching text from the taxonomy list.
- Applied a similarity threshold of 40% — pairs below this score were not tagged.
- The 40% limit helped include slightly misspelled or incomplete text while avoiding wrong matches.

Step 3 – Logical Mapping

- Root Cause → Matched from the Cause column.
- Symptom Condition → Derived from the Complaint column.
- Symptom Component → Derived from the Complaint column.
- Fix Condition → Mapped using the Correction column.
- Fix Component → Mapped using the Correction column.

Step 4 – Output Generation

- Final tagged data was saved into a new Excel file named Tagged_Task_Output.xlsx.
- This allowed easy verification and further analysis.

Insights and Food for Thought

Frequent Root Causes

- Help identify recurring production or design problems.
- Useful for improving process efficiency and product quality

Repeated Symptom Conditions

- Indicate common failure patterns or weak points in product performance.
- Can guide teams to focus on preventive measures in those areas.

Fix Component Analysis:

- Highlights which parts or processes require frequent maintenance.
- Helps in resource planning and inventory optimization.

Automation Benefits:

- The tagging process reduces manual effort and human error.
- Creates a structured dataset that can be easily analyzed.

Future Applications:

- Enables data-driven root cause analysis and continuous improvement.
- Can be extended for predictive maintenance models and interactive dashboards.
- Supports early detection of issues, improving reliability and customer satisfaction.

Thank You