TASK 1: DATA TAGGING

ORIGINAL EXCEL DATA:

Primary Key	Order Date	roduct Categor	Complaint	Cause	Correction	Root Cause	Symptom Condition	Symptom Component	Symptom Condition	Symptom Component	Symptom Condition	Symptom Component	Fix Condition	Fix Component	Fix Condition	Fix Component	Fix Condition	Fix Component
SO0026296-1	08-03-2023	SPRAYS	cab on P clips and air ducting left	Not tighten at factory.	ALL P CLIPS, NUTS, AND BOLTS	Not Tightened	Loose	Cab P Clip	Loose	Left-Air Duct	Loose	Bulkhead Connector	Retightened	Cab P Clip	Retightened	Left Air Duct	Retightened	Bulkhead Connecto
SO0026385-1	08-03-2023	SPRAYS	Fuel door will not stay open	ANYWHERE ON MACHINE	INSTALLED OR ANYWHERE ON	Not Installed	Won't stay open	Fuel Door					Installed	Gas Strut				
SO0026385-11	08-03-2023	SPRAYS	braided steel, crushed	braided steel, crushed when	SYSTEM.REMOVE ASSOCIATED													
SO0028352-1	08-03-2023	SPRAYS	machine	SWIVEL FITTING LEFT LOOSE AT	SWIVEL FITTING LEFT LOOSE AT													
SO0028770-1	08-03-2023	SPRAYS	UNLOCKS.	UNLOCKS WERE NOT INSTALLED.	RAN AND TESTED.													
SO0029596-1	08-03-2023	SPRAYS	RETURN LINE TO PRODUCT PUMP	Coupler was leaking.	WITH HAMMER AND SOCKET IN													
SO0058466-2	05-05-2023	SPRAYS	BOOM TO MOUNT SMV SIGN	NOT INCLUDED FROM FACTORY	MISSING BRACKETS AND BOLTS													
SO0058466-3	16-05-2023	SPRAYS	MACHINEPICTURES INCLUDED	CONNECT STICKING OUT OF	ON MALE QUICK CONNECT													
SO0058466-4	16-05-2023	SPRAYS	OIL LEAK	BLOWN ORING	OUT TO SPRAYER GOT TO													
SO0058466-5	16-05-2023	SPRAYS	HARNESS BROKE	POOR MATERIAL IN HARNESS	REPLACED THE NCV HARNESS													
SO0058668-1	31-05-2023	SPRAYS	RINSE TANK LEAK	SIGHT GLASS TUBE IS LEAKING	181342MILEAGE OUT:													
SO0058674-1	31-05-2023	SPRAYS	FUEL SENDER OPEN	51457614UNIT HRS: 19.75	FUEL SENDER CODE. PULLED													
SO0058674-2	31-05-2023	SPRAYS	HYDRAULIC LEAK	COOLER FITTINGUNIT HRS 21PN:	LEAK. FOUND IT WAS DRIPPING													
SO0058727-15	19-06-2023	SPRAYS	HYDRAULIC LEAK	HOSE, PART NUMBER 90391635	TO FIND LEAKING HOSE.FOUND													
SO0058727-2	05-06-2023	SPRAYS	HITTING CRADLE	SENSORS OUT OF RANGE	to go over the PDI of the													
SO0058727-5	09-06-2023	SPRAYS	CODE: F ¼Â° 354›ÀÂ°31.	DRIP DOWN CONTAMINATING	SENSOR.LUBRICANT USED ON O													
SO0058727-6	09-06-2023	SPRAYS	AUTO BOOM WON'T WORK	SENSORAUTOBOOM NODE	OUT: 17263MILEAGE IN:													
SO0058796-1	13-07-2023	SPRAYS	DEF ERROR CODES	MODULE	HOOKED UP EST AND FOUND													
SO0058796-4	17-07-2023	BALER	Massive product leak under machine	TRANSDUCER WAS ONLY SCREWED IN A THREAD	FOUND ELBOW FITTING FOR PUMP. TRANSDUCER WAS ONLY SCREWED IN A THREAD, RESULTING IN BROKEN ELBOW AND PULLING THREAD OUT OF PUMP HOUSING.REPAIR AND RE-TAP THREADS IN PUMP HOUSING.INSTALL PIPE FITTING UNTIL NEW PART ARRIVED.SECOND TRIPREMOVE AND REPLACE ELB													
SO0058796-4	17-07-2023	BALER	Condenser loose.	Faulty bolts in bracketPN: 47953465Unit Hrs: 265	Drove machine into shop. Found that the condenser bracket had broken. Hook ac machine to unit. Recover ac system. Remove condenser/bracket. Went to install new condenser/bracket and found it was wrong. Customer needed machine. Tapped new larger threads in													

AFTER DATA TAGGING:

- Task1Data Tagging
- Python file

SUMMARY REPORT:

Tagging Complaints Data using LLMs-Large language models (Here I've used GPT-4)

To automate structured tagging of complaint logs I have utilized Openais GPT-4 to extract five key diagnosis fields: Root Cause, Symptom condition, System Component, Fix Condition, and Fix Component. Every unstructured complaintt record contained a description of the issue-Complaint its diagnosed reason-cause and the solution performed-Correction. The tagging was handled using python and openai api. The script processed each row by sending a structured prompt to gpt-4 asking to extract. A root cause(Not installed, Factory didn't tighten) Up to 3 Symptom conditions and components like loose, cab p clip. And 3 fix conditions and components Retightened, left duct.

The prompt was designed to guide LLM to output in a format ready to be mapped into a dataframe. Any missing symptom or fix was left blank to maintain data integrity and avoid false positives. The response was parsed row-by-row and populated into a structured CSV format.

Once Structured it helps unlock recurring root cause as frequently occurring issues for example loose fittings, missing parts can highlight manufacturing or weak design. Grouping symptom conditions can help flag indicators of failures, improving preventive maintainence models. Tracking which fixes work best for each issue enables building a knowledge base for field technicians. Structured cause data can help segregate factor origin vs usage origin issues, useful for warranty claims and cost sharing.

PS:

Issues like: Loose p clips, oil leaks from fittings, missing gas struts. These are captured in your structured data under Symtom Conditions and Symptom Component. If you track how often and where these symptoms happens over time we can spot which machnine or parts break down before they fail.Plan maintenance in advance reducing breakdowns and costly delays. Using past symtoms patterns to predict and prevent future failures.

If Root cause data shows that problems may arise like: Not tightened, Not installed..then problem isn't in that field its at the factory meaning, the factory quality checks are failing. The manufacturer needs to update their assembly process train staff better and improve inspection.

***

~ Diya Haldankar