Task 2:
Data Analysis and Insights Generation

A. Column Analysis

The dataset contains identifier fields, categorical attributes, numeric measures, date columns, and free-text descriptions. Identifier columns help track vehicles and service transactions across the system. Categorical columns describe vehicle configuration, manufacturing details, dealer information, and parts involved. Numeric columns support quantitative analysis such as frequency and impact of issues. Date columns enable trend and time-based analysis. Free-text fields provide detailed customer complaints and technician repair notes, adding valuable context to the data.

B. Data Cleaning Summary

Data cleaning was performed to improve consistency and reliability of the dataset. Missing values were handled based on data type, using placeholders for categorical fields and median values for numeric fields. Text inconsistencies such as extra spaces and inconsistent capitalization were standardized. Numeric columns were validated to ensure correct data types. Duplicate records were checked and removed where applicable. These steps ensured the data was clean and ready for analysis.

C. Visualizations

Visualizations were created to highlight meaningful patterns in the data. Bar charts were used to compare issue distribution across plants, platforms, body styles, and components. These charts make it easy to identify categories with higher issue concentration. A line chart was used to visualize issues over time using repair dates. This helped identify trends

and potential spikes in failures. The visualizations support quick and clear interpretation for stakeholders.

D. Generated Tags & Key Takeaways

Tags were generated from free-text fields to summarize failure conditions, affected components, and repair actions. This converted unstructured text into structured and analyzable information. The analysis shows frequent leak-related and sensor-related issues across records. A small number of components account for a large share of failures. Most repairs involve replacement or tightening actions. These insights suggest the need for improved assembly checks and standardized repair practices.