# Task – 2

# Summary and Insights

## Summary of Generated Tags

The generated tags primarily focus on key failure conditions, components, and corrective actions related to vehicle repairs. Some of the most frequent tags include: Steering, Wheel, Heated, Module, Replace, Horn, Test, Checked, Cruise. These tags indicate common issues reported by customers and the actions taken to resolve them. The dataset suggests that steering wheel-related problems, heating malfunctions, and module replacements are frequent service concerns.

## Potential Insights Derived

1. High Incidence of Steering Wheel Issues – A significant portion of the dataset revolves around steering wheel failures, indicating a potential design or manufacturing defect.  
2. Frequent Module and Component Replacements – The presence of terms like 'module,' 'replace,' and 'checked' suggests that certain components may have recurring failures requiring replacement.  
3. Customer Concerns vs. Actual Fixes – Comparing 'CUSTOMER\_VERBATIM' with 'CORRECTION\_VERBATIM' reveals that some issues reported by customers may have different underlying causes than initially perceived.  
4. Impact on Service Efficiency – The frequent occurrence of 'test' and 'checked' implies extensive diagnostic procedures, potentially affecting service time and efficiency.

## Actionable Recommendations for Stakeholders

• Quality Control Improvement: Conduct a detailed analysis of the steering wheel components to identify potential manufacturing defects.  
• Preventive Maintenance Guidelines: Introduce predictive maintenance strategies for components like steering modules and heated elements to reduce frequent failures.  
• Enhanced Customer Communication: Develop better diagnostic procedures to align customer-reported issues with actual technical failures for better service accuracy.  
• Training for Technicians: Since 'checked' and 'test' frequently appear, enhancing technician training can improve repair efficiency and reduce unnecessary part replacements.  
• Feedback Mechanism: Implement a structured system to track customer-reported issues against actual fixes to improve customer satisfaction and product design.

## Data Discrepancies and Approach to Handling Them

1. Missing Values: Some records had null values in 'CUSTOMER\_VERBATIM' and 'CORRECTION\_VERBATIM.' These were handled by filling missing values with blank spaces before processing.  
2. Duplicate Records: Potential duplicate entries were checked based on 'VIN' and 'TRANSACTION\_ID' to ensure unique repair transactions.  
3. Unstructured Text Processing: Text fields contained variations of similar issues (e.g., 'heated steering wheel not working' vs. 'heated wheel inop'). Standardization was done using text preprocessing techniques.  
4. Irrelevant Words Removal: Common stopwords (e.g., 'the,' 'is,' 'was') were filtered out using TF-IDF and frequency-based methods to extract meaningful keywords.