Report: Vehicle Repair and Issue Insights

1. Column Analysis

The dataset captures **customer complaints, repair information, and associated costs**. Key columns include:

Column	Description	Observations
CUSTOMER_VERBATIM	Customer complaint description	Textual; required cleaning and translation for non-English entries
CORRECTION_VERBATIM	Description of corrective action	Combined with customer verbatim for keyword extraction
COMPLAINT_CD	Complaint code	Categorical; used to group and analyze repair age
TOTALCOST	Total repair cost	Missing/negative values imputed with median
LBRCOST	Labor cost	Used to evaluate dealer cost contribution
DEALER_NAME	Dealer performing repair	Used for cost and labor distribution analysis
COUNTRY_SALE_ISO	Country code of sale	Used for regional analysis

Observation: Several columns contained nulls, invalid entries, or duplicates. Example: TRANSMISSION_TRACE_NBR was partially empty and replaced with placeholders.

2. Data Cleaning Summary

Key steps:

1. **Dropped irrelevant columns** (e.g., campaign numbers, non-causal parts).

2. **Numerical cleaning:** Negative/missing values in KM, TOTALCOST, REPAIR_AGE, and LBRCOST replaced with medians.

3. **Categorical cleaning:** Missing values in DEALER_NAME, PLANT, STATE filled with mode.

4. **Text standardization:** Converted to uppercase; removed leading/trailing spaces.

5. **Translation:** Non-English complaints translated into English.

6. **Keyword tagging:** Combined CUSTOMER_VERBATIM and CORRECTION_VERBATIM for issue extraction.

Discrepancies:

- Missing complaint codes (primary keys) , incomplete textual description and blanks.

3. Visualizations (Top 5 values)

1. **Top Keywords:** Frequent terms include *STEERING* and *HEATER*.

2. **Avg Repair Age by Complaint Code:** Highlights complaints resolved slower vs. faster.

3. **Avg KM by Repair Age:** Shows vehicle usage patterns vs. time-to-repair.

4. **Issue Category Distribution:** Steering and electrical dominate issue categories.

5. **Country-wise Sales & Labor cost by dealer:** Top 5 countries/dealer by sales share.

4. Generated Tags & Insights

- **Tags created** from combined customer and corrective text.

- **Top issue categories:** Steering, Fabric, Heating.

Key Insights:

- Steering complaints are most frequent → potential design/component issues.

- Repair age rises with higher KM → opportunity for preventive maintenance.

- Few dealers drive majority of labor costs → efficiency gaps exist.

5. Actionable Recommendations

1. **Product Improvement:** Address recurring steering/heating issues with design fixes.

2. **Preventive Maintenance:** Implement KM-based maintenance reminders to reduce delay.

3. **Dealer Management:** Audit high-cost dealers to optimize labor usage.

4. **Regional Focus:** Prioritize quality assurance in high-complaint countries.

Conclusion

The integration of complaint narratives with operational metrics highlights **where failures occur, how much they cost, and which dealers or regions are most affected**. By acting on these insights, stakeholders can improve customer satisfaction, reduce repair costs, and strengthen long-term product reliability.