Repair Data Analysis Report

### ****Dataset Description****

1. **VIN (Vehicle Identification Number)**: A unique code assigned to each vehicle, which helps in identifying specific vehicles. It’s like a fingerprint for cars, used to track the vehicle’s history and specifications.
2. **TRANSACTION\_ID**: This is a unique identifier for each repair or service transaction. It helps to track each repair event in the dataset.
3. **CORRECTION\_VERBATIM**: Detailed description of the repair or correction performed. This column provides a free-text explanation of what was fixed or corrected in the vehicle.
4. **CUSTOMER\_VERBATIM**: This column records the customer's complaint or issue with the vehicle, written in their own words.
5. **REPAIR\_DATE**: The date when the repair or service was completed. This helps in understanding the timeline of repairs.
6. **CAUSAL\_PART\_NM**: This column contains the name of the part that caused the issue, leading to the repair. It identifies the root cause of the problem.
7. **GLOBAL\_LABOR\_CODE\_DESCRIPTION**: A description of the labor operation performed during the repair, based on a global coding system. It gives details about the type of labor involved.
8. **PLATFORM**: The vehicle’s underlying design or architecture. It indicates the base design that the vehicle is built on, such as the type of vehicle platform.
9. **BODY\_STYLE**: The type or design of the vehicle’s body. For example, this could be sedan, SUV, truck, etc.
10. **VPPC (Vehicle Product Program Code)**: A code used to identify the vehicle in a product program. This is used for tracking specific vehicle models and configurations.
11. **PLANT**: The manufacturing plant where the vehicle was assembled. It helps in tracing the vehicle's production history.
12. **BUILD\_COUNTRY**: The country in which the vehicle was manufactured. This gives geographical information about where the vehicle was produced.
13. **LAST\_KNOWN\_DLR\_NAME**: The name of the last dealer who interacted with or serviced the vehicle.
14. **LAST\_KNOWN\_DLR\_CITY**: The city where the last known dealer is located. It helps in identifying where the vehicle was last serviced.
15. **REPAIRING\_DEALER\_CODE**: A code identifying the dealer that performed the most recent repair or service on the vehicle.
16. **DEALER\_NAME**: The name of the dealer performing the current repair.
17. **REPAIR\_DLR\_CITY**: The city where the repair dealer is located.
18. **STATE**: The state in which the repair dealer is located.
19. **DEALER\_REGION**: The region associated with the dealer performing the repair. It helps in understanding the geographical service area.
20. **REPAIR\_DLR\_POSTAL\_CD**: The postal code of the repair dealer’s location.
21. **REPAIR\_AGE**: The age of the vehicle at the time of repair, usually expressed in months or years. This can indicate whether older vehicles are more likely to need repairs.
22. **KM**: The distance traveled by the vehicle, recorded in kilometers. It helps in tracking how much the vehicle has been used and whether higher mileage is associated with more repairs.
23. **COMPLAINT\_CD\_CSI**: Customer Satisfaction Index code for the complaint. It categorizes the complaints based on their severity or nature.
24. **COMPLAINT\_CD**: A code that categorizes the type of complaint made by the customer. It helps in analyzing the most common issues vehicles face.
25. **VEH\_TEST\_GRP**: This refers to the vehicle test group, which could indicate the vehicle’s emissions or compliance with specific vehicle regulations.
26. **COUNTRY\_SALE\_ISO**: The ISO code of the country where the vehicle was sold. This helps to identify the geographical sales market for the vehicle.
27. **ORD\_SELLING\_SRC\_CD**: A code representing the source from which the vehicle order or sale originated.
28. **OPTN\_FAMLY\_CERTIFICATION**: Certification information related to the vehicle’s option families. Option families refer to the different configuration options available for the vehicle.
29. **OPTF\_FAMLY\_EMISSIOF\_SYSTEM**: This column indicates the emission system associated with the vehicle’s option family.
30. **GLOBAL\_LABOR\_CODE**: A global labor operation code that identifies specific labor tasks performed during repairs.
31. **TRANSACTION\_CATEGORY**: This categorizes the transaction, such as warranty, service, or repair. It helps in understanding the nature of the transaction.
32. **CAMPAIGN\_NBR**: The number of the associated service campaign or recall. This column is useful for identifying vehicles that are part of a recall or repair campaign.
33. **REPORTING\_COST**: The cost reported for the repair or service performed. It reflects the expense associated with the repair.
34. **TOTALCOST**: The total cost incurred for the repair, including parts and labor. It helps in understanding the financial aspect of repairs.
35. **LBRCOST**: The labor cost incurred during the repair. This is part of the total cost of the repair.
36. **ENGINE**: The engine code or identifier. It helps in identifying the specific engine model in the vehicle.
37. **ENGINE\_DESC**: A description of the engine type or specifications.
38. **TRANSMISSION**: The transmission code or identifier for the vehicle’s transmission system.
39. **TRANSMISSION\_DESC**: A description of the transmission system in the vehicle.
40. **ENGINE\_SOURCE\_PLANT**: The plant where the engine was manufactured.
41. **ENGINE\_TRACE\_NBR**: A traceability number for the engine, used for tracking its history.
42. **TRANSMISSION\_SOURCE\_PLANT**: The plant where the transmission was manufactured.
43. **TRANSMISSION\_TRACE\_NBR**: A traceability number for the transmission system.
44. **SRC\_TXN\_ID**: The source transaction ID, which links the current transaction to the original record.
45. **SRC\_VER\_NBR**: The version number of the source transaction record. This helps in tracking different versions of a transaction record.
46. **TRANSACTION\_CNTR**: A counter for the number of transactions associated with a specific vehicle. This helps in understanding how many times a vehicle has been serviced or repaired.
47. **MEDIA\_FLAG**: An indicator for whether media (images, videos) is attached to the transaction record. This can be useful in visual documentation of repairs.
48. **VIN\_MODL\_DESGTR**: This column links the VIN to the model designation, indicating which vehicle model is associated with the VIN.
49. **LINE\_SERIES**: This refers to the series or product line of the vehicle, which helps in categorizing the vehicle based on its model range.
50. **LAST\_KNOWN\_DELVRY\_TYPE\_CD**: A code that indicates the type of delivery used for the last known delivery of the vehicle.
51. **NON\_CAUSAL\_PART\_QTY**: The quantity of parts used in the repair that were not identified as the causal part. This helps in understanding what additional parts were replaced during the repair process.
52. **SALES\_REGION\_CODE**: A code representing the sales region of the vehicle, which helps in understanding the geographic area in which the vehicle was sold.

This report provides an analysis of the vehicle repair data. It covers various aspects including repair costs, failure parts, dealer performance, and platform usage. The visualizations and analysis aim to identify trends and insights that can help improve the repair processes.

# 1. Total Repair Costs

The distribution of repair costs shows a wide range, with most repairs having lower costs. The histogram below helps visualize how repair costs are spread out. This can help identify if there are areas where costs are unusually high, which might need further investigation.

Visual: Distribution of Total Repair Costs

# 2. Correlation between Key Variables

A heatmap was used to visualize the correlation between variables such as 'KM', 'TOTALCOST', 'REPAIR\_AGE', 'REPORTING\_COST', and 'LBRCOST'. This helps to understand how different factors relate to each other, especially how repair age and mileage affect the total cost of repairs.

Visual: Correlation Heatmap

# 3. Top Failure Parts

The top 10 most frequent failure parts were identified. This information is useful for understanding which parts are most prone to issues and can help improve parts management and inventory control.

Visual: Top 10 Most Frequent Failure Parts

# 4. Repairs by Vehicle Test Group

This analysis shows which vehicle test groups have the most repairs. It helps to identify patterns in which vehicle types experience the most issues. This can help prioritize repairs and resources for specific vehicle categories.

Visual: Top 5 Counts of Repairs by Vehicle Test Group

# 5. Repairs by State

The repair costs across different states were analyzed. This shows the total repair cost per state and helps identify regions where repairs are most frequent or costly. It could be useful for targeting resources to specific areas.

Visual: Top 10 Total Repair Costs by State

# 6. Top Dealers by Number of Repairs

The dealer postal codes were analyzed to find the top 10 dealers with the most repairs. This helps in evaluating dealer performance and identifying areas where repairs are concentrated.

Visual: Top 10 Dealers Postal Code by Number of Repairs

# 7. Repairs by Platform

The number of repairs per platform was analyzed. This shows how repairs are distributed across different platforms, helping to identify if certain platforms experience more frequent issues than others.

Visual: Number of Repairs by Platform

# 8. Labor Codes Frequency

A pie chart shows the frequency of different labor codes used for repairs. This helps understand which labor codes are most commonly associated with repairs.

Visual: Labor Codes Frequency

# 9. Repairs by Body Style

The number of repairs was analyzed by vehicle body style. This helps identify which body styles are most prone to repair issues, which can help manufacturers or service providers focus on particular vehicle designs.

Visual: Number of Repairs by Body Style

### ****Top 5 Critical Columns for Stakeholders****

1. **TOTALCOST**:
   * **Reasoning**: This column represents the total cost incurred for the repair, including parts and labor. Understanding the cost trends is crucial for stakeholders as it directly impacts the business's profitability and helps in identifying areas where cost optimization is needed.
2. **REPAIR\_AGE**:
   * **Reasoning**: The age of the vehicle at the time of repair is essential for identifying patterns in repair frequency and cost. Older vehicles might require more repairs, and this information helps in assessing how repair strategies should differ for different vehicle ages.
3. **CAUSAL\_PART\_NM**:
   * **Reasoning**: The name of the part that caused the repair is essential for identifying recurring problems and part failure rates. Understanding which parts are most prone to failure can help in improving vehicle design, parts inventory management, and reducing repair frequency.
4. **STATE**:
   * **Reasoning**: The geographical location of the repair gives insight into regional differences in repair needs. It helps in understanding whether certain regions experience more repairs due to environmental factors, local road conditions, or regional vehicle issues. This information is valuable for targeting specific regions for marketing, parts supply, and customer service.
5. **PLATFORM**:
   * **Reasoning**: The platform on which the repair service was conducted is important for understanding the efficiency and availability of different service channels. Knowing which platforms are being used more can guide improvements in service offerings and customer experience across different platforms.

# Conclusion

This analysis provides valuable insights into vehicle repair trends, highlighting the most common failure parts, the vehicle categories that require the most repairs, and the geographical areas where repairs are concentrated. These findings can help optimize repair operations, improve resource allocation, and focus on high-need areas.