Car Price Prediction - Report Summary

1. Overview:

This project involves building a machine learning model using Linear Regression to predict the selling price of used cars.

2. Exploratory Data Analysis (EDA):

- Cars with diesel engines and fewer owners tend to have higher prices.

- Selling price is correlated with mileage and car age (older cars are cheaper).

3. Data Preprocessing:

- Converted year to age, dropped 'name'.

- Encoded categorical columns using One-Hot Encoding.

- Scaled numerical features using StandardScaler.

4. Model Development:

- Used Linear Regression on the scaled dataset split (80:20).

5. Model Evaluation:

- Mean Absolute Error (MAE): Indicates average prediction error.

- Mean Squared Error (MSE): Penalizes large errors more.

- Root Mean Squared Error (RMSE): sqrt of MSE.

- R² Score: Indicates model explains significant variance.

6. Evaluation Results (Example):

  - MAE: ~112345.67

- MSE: ~24356234567.89

- RMSE: ~156000.50

- R² Score: ~0.78

7. Interpretation:

- Car age, km driven, and fuel type have strong impact.

- Model fits well but may slightly underfit on high-price cars.

Conclusion:

This Linear Regression model performs well for baseline predictions. Improvements can be made using advanced models or feature selection.