Key Insights:

1.) Vehicle Age: Newer cars command higher prices

2.) Odometer: Lower mileage is valued more

3.) Manufacturer: Brands matter significantly (e.g., Toyota, BMW)

\* Random Forest performed best (lowest RMSE, highest robustness)

\* Model related:

1.) Cross-validated for reliability

2.) Random Forest had the least RMSE

3.) Lasso faced convergence issues due to complexity of categorical features

Value for the Business:

• Use model insights to focus inventory acquisition on high-value vehicle profiles

• Support smarter pricing with data-backed feature importance

Modeling results in detail:

Summary based on each model is as shown below

1.) Linear regression: (RMSE =0.007502)

Equation: please the code output

Key features: Top 3 key features are:

manufacturer\_morgan 2.160091

manufacturer\_ferrari 1.543470

manufacturer\_datsun 0.951816

2.) Ridge regression: (RMSE =0.007495)

Equation: please the code output

Key features: Top 3 key features are:

manufacturer\_ferrari 1.497569

manufacturer\_datsun 0.880217

manufacturer\_mercury 0.670043

3.) Lasso regression: (RMSE = 0.006336)

Equation: price = 10.34 + -0.0844\*log\_odometer + 0.1574\*cylinders\_scaled + -0.2404\*vehicle\_age\_scaled

Key features: Top 3 key features are:

vehicle\_age\_scaled 0.240411

cylinders\_scaled 0.157411

log\_odometer 0.084429

All others are forced to zero

4.) RandomForest regression: (RMSE = 0.004827)

Equation:See the code output

Key features: Top 3 key features are:

vehicle\_age\_scaled 0.240411

cylinders\_scaled 0.157411

log\_odometer 0.084429

All others are forced to zero

5.) Gradient Boosting Regressor Model: RMSE 0.005580

Equation:not vaiable as it is a non-linear ensemble of decision trees

Key features: Top 3 key features are:

vehicle\_age\_scaled 0.466033

log\_odometer 0.163960

cylinders\_scaled 0.144242

All others are NOT forced to zero

Based on the best model i.e RandomForest the key features are vehicle\_age\_scaled,cylinders\_scaled and log\_odometer.