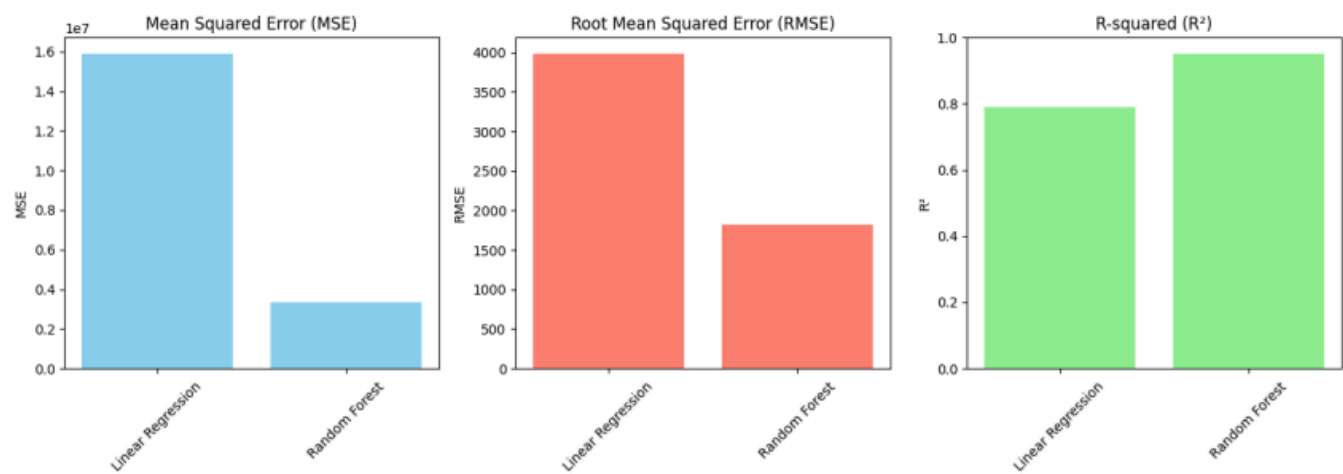**Visualization for MSE, RMSE and R2**

Model Comparison Report:

Predicting Car Prices for Geely Auto's Market Entry Strategy Objective To develop a predictive model for car prices based on multiple independent variables. This model will help Geely Auto understand pricing dynamics in the US market and guide business and design strategy. Models Evaluated 1. Linear Regression 2. Random Forest Regression (with tuned n_estimators, max_depth, min_samples_split, and min_samples_leaf).

Performance Metrics Model | Best Parameters | MSE | RMSE | R²

| ————————————————————-- |

Linear Regression |Default |

15916389 | 3989 | 0.79 |

Random Forest | n_estimators=200, max_depth=30, min_samples_split=2, in_samples_leaf=1 | 3326665 | 1823 | 0.95 |

Model Insights

1. **Linear Regression**:

   • Baseline model with satisfactory performance.

   • R² of 0.79 indicates it captures most of the variation but has limited flexibility with non-linear relationships.

2. **Random Forest Regression**:

   • Non-linear model with ensemble learning, aggregating multiple decision trees.

   • Outperforms other models with an R² of 0.95, capturing more complex relationships in the data and reducing MSE to 3326665, making it the most accurate for car price

prediction.

Recommendation for Production

Based on the evaluation metrics, Random Forest Regression is the best model for production deployment. It offers the highest accuracy, capturing the non-linear relationships between features and car prices more effectively than linear models. Random Forest's flexibility and robustness in handling varied data types and interactions make it suitable for predicting car prices in the competitive and dynamic US market.