**Model Evaluation**

Random Forest Regressor Model: -

Random forest regressor model is performing well with our dataset, with a high R² score suggesting that it explains most of the variability in the fare prices. The RMSE provides an error measure in the same unit as the fare, showing the average deviation of the predicted fares from the actual fares.

Xg-Boost Model: -

Xg-boost model performing well with predicting flight prices. It also demonstrates strong performance with high R² and relatively low MAE and RMSE values. It indicates that the model is highly effective at predicting flight fares, with the majority of the variance in the data being explained by the model and the average prediction errors being reasonably small. This means the model can be trusted to make accurate predictions most of the time.

Linear Regression Model: -

Linear regression model is not performing as well as the Xg-boost model. It has higher errors, meaning its predictions are less accurate. While linear regression is simpler and easier to interpret, it does not capture the complexity of the data as effectively as the more sophisticated XGBoost model.

Stats Model: -

On stats model for flight price prediction, We are choosing Adj R-squared over R-squared because When comparing models with different numbers of predictors, Adjusted R-squared provides a fairer comparison by accounting for the number of predictors. Moreover, Avoid Overfitting with Adjusted R-squared helps in identifying overfitting by penalizing the inclusion of irrelevant predictors. Also provides model accuracy it measure of the goodness of fit, particularly in large datasets with many predictors. choosing Adjusted R-squared over R-squared ensures that the model evaluation considers the number of predictors, leading to more reliable and meaningful interpretations of the model's performance.

In the context of our regression model, p-values lower than 0.05 are crucial as they indicate that the predictors are statistically significant, meaning there is a less than 5% probability that the relationship between the predictor and the dependent variable (Fare) is due to chance. This significance helps in validating that the predictors are genuinely contributing to explaining the variance in the Fare, thereby enhancing the model's accuracy and robustness. By focusing on predictors with p-values below this threshold, we can streamline the model, eliminate noise from non-significant variables, and avoid overfitting, leading to better performance and interpretability of the model on new data.