Training Analysis Report

This report summarizes the training analysis performed in this notebook, focusing on the models explored and their performance during the training and validation phases.

1. Model Exploration:

Several regression models were explored to predict the 'Overall Score Excluding Monetary Freedom':

- Linear Regression
- Ridge Regression
- Lasso Regression
- Polynomial Regression
- Decision Tree Regressor
- Random Forest Regressor
- Gradient Boosting Regressor
- XGBoost Regressor

2. Training Process and Validation:

- Models were trained on the train_df_imputed dataset, which contains data from 1995 to 2019 after cleaning and imputation.
- **Time Series Cross-Validation (TSCV)** was used during the initial model exploration (specifically for the Linear Regression pipeline) to evaluate performance on different time folds within the training data. This helps assess the model's stability and generalization capability over time before evaluating on the independent test set.
  - The average MSE across the 5 folds of TSCV for the initial Linear Regression pipeline was **{{average_mse:.4f}}**.
- Individual tree-based models (like the Decision Tree Regressor) were evaluated using cross-validation on the training set to tune hyperparameters (e.g., max_depth).
  - Cross-validation for the Decision Tree Regressor with varying max_depth showed decreasing MSE as depth increased, with the best average CV MSE of **{{lowest_avg_mse:.4f}}** achieved at a max_depth of **{{best_max_depth}}**.

3. Key Observations from Training:

- The cross-validation results provided insights into how the models performed on different subsets of the training data over time.
- The performance metrics obtained during cross-validation helped in comparing the initial models before the final evaluation on the independent test set.

- Tuning hyperparameters using cross-validation (as shown with the Decision Tree) is a crucial step in optimizing model performance and avoiding overfitting on the training data.

4. Transition to Test Set Evaluation:

While cross-validation provides a good estimate of performance within the training distribution, the true measure of generalization is the evaluation on the independent test set (data from 2020 onwards). The results from the test set evaluation (detailed in the Model Evaluation & Optimization section) were used for the final model selection.

This training analysis provides a foundation for understanding how the models were developed and initially assessed before their final performance was determined on the unseen test data.