Proposed Methodology

We applied multiple regression models to predict **G3** using the dataset from the Irvine repository. Our goal was to identify the most suitable model by balancing underfitting and overfitting while maximizing predictive accuracy. Below are the methodologies and results:

1. Linear Regression

We began with linear regression and enhanced it with techniques such as gradient boosting and L1/L2 regularization. However, L1 and L2 regularization were found to over-penalize the model, limiting its effectiveness, especially since regular linear regression was not prone to overfitting in the first place.

- **Cross-Validated Mean Squared Error (MSE)**: 11.93
- **Mean Squared Error (MSE)**: 14.44
- **R²**: 0.066

Linear regression showed limited predictive power, with an **R²** close to zero, indicating it could not capture the complex relationships between the features and **G3**.

2. Polynomial Regression

Polynomial regression was employed to capture non-linear patterns in the data. A degree of 2 (Poly2) emerged as the best-performing model. While it slightly underfits the data, it provided a reasonable balance between bias and variance. However, polynomial regression with a degree of 3 (Poly3) tended to overfit, capturing noise rather than meaningful patterns.

- **Polynomial Regression (Degree 2)**:
  - Balanced predictive performance with slight underfitting.
- **Polynomial Regression (Degree 3)**:
  - Overfitted the data, leading to reduced generalizability.

3. Decision Tree Regression

Decision trees were also tested, but their performance fell short compared to the Poly2 model. Decision trees tended to overfit the small dataset or underperform with limited complexity. Below are the results for different depths:

Depth	MSE	R²	Regularized Risk
3	4.70	0.77	5.50
4	6.20	0.70	7.80
5	6.09	0.70	9.18
6	4.27	0.79	9.17
7	4.30	0.79	11.99

Although decision trees performed well in capturing non-linear patterns, their tendency to overfit or underperform compared to Poly2 makes them a less suitable choice.

Best Model: Polynomial Regression (Degree 2)

The **Polynomial Regression (Degree 2)** model emerged as the most appropriate model for this dataset, striking a balance between bias and variance. It demonstrated strong predictive capabilities while avoiding the pitfalls of overfitting or underfitting. Future work could explore ensemble methods or hybrid approaches to further refine the predictions and enhance model robustness.