Decoding Student Success:
An Analysis of Academic Performance

Name: Eric
Project Track: Data Science
Major: Computer Science & Engineering

Problem Statement:

Student academic performance is influenced by a wide range of factors, including study habits, socioeconomic background, attendance, and personal well-being. However, these relationships are often unclear. This project aims to use data-driven analysis to uncover the key determinants of student success and to understand how different features contribute to variations in academic outcomes.

Dataset:

The dataset was obtained from Kaggle and contains information from 200 students, including study hours, sleep duration, attendance rate, previous scores, and final exam scores.

Steps Taken:

1. Loaded and cleaned the dataset using Pandas.
2. Used Matplotlib, Seaborn, and Plotly to visualize feature distributions and basic relationships.
3. Built Linear Regression and Random Forest models to analyze which factors best predict exam performance.
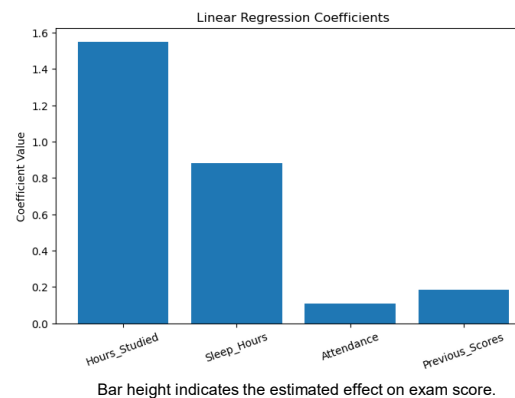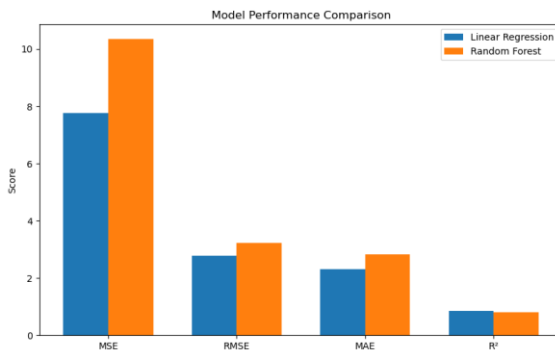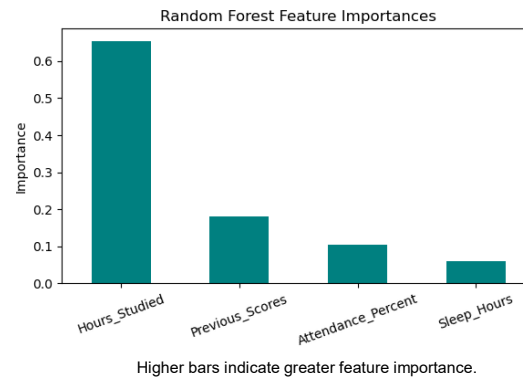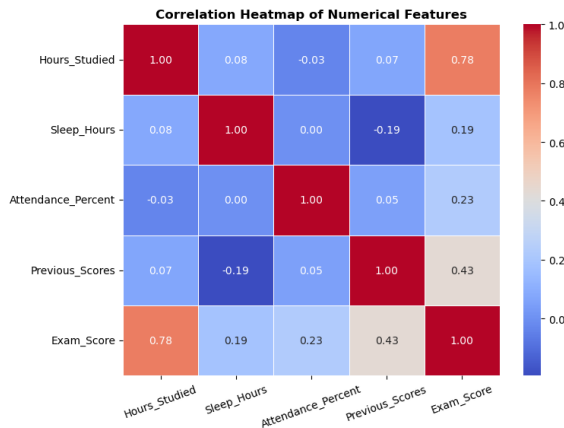
Modeling:

Two models were applied — Linear Regression and Random Forest — to predict exam scores. Their performance was compared using MSE, RMSE, MAE, and R^2.

Correlation Heatmap of Numerical Features

Random Forest Feature Importances
Higher bars indicate greater feature importance.

Model Performance Comparison

Linear Regression Coefficients
Bar height indicates the estimated effect on exam score.

Results:

The analysis shows that hours studied and previous scores are the strongest predictors of exam performance. Linear Regression results indicate that each additional hour of study is associated with roughly a **1.55-point increase** in exam score, while sleep contributes about **0.90 points**. Random Forest feature importances also rank study hours as the most influential factor.

An important insight is that **consistent study habits** matter more than extreme behaviors—students do not need to stay up late, as losing sleep harms health and does **little** to improve scores. While the model performs well, **exam difficulty and other external factors** may also affect outcomes and could be explored in future research. Overall, the findings suggest that balanced routines and steady studying are key to success, offering insights that can help schools support healthier and more effective academic strategies.

BDAA
BIG DATA & ANALYTICS ASSOCIATION