INTRODUCTION OF AI IN EDUCATION

SIDDHI JALAN
SAKSHI PRIYA
ANSHITA VERMA

PROBLEM STATEMENT

Student Performance Analysis and Prediction (20% of Focus)

GOAL   Analyze student performance data to identify patterns and predict future performance

ANALYSIS

KEYWORDS

- Perform correlation analysis to identify relationships between performance in different subjects.

- Explore regression models to predict student performance

REGRESSION ANALYSIS

CORRELATION ANALYSIS

PROJECT INTRODUCTION

This project delves into leveraging AI to enhance education by analyzing student performance and predicting academic trends. It aims to transform teaching strategies by harnessing groundbreaking insights from student data, ultimately improving learning experiences and educational outcomes.

PROJECT TIMELINE

Regression Analysis

In our regression analysis, we explored linear regression, random forest regression, support vector regression (SVR), and decision tree regression models to understand relationships between variables and predict outcomes.

Results and Analysis

Results and analysis entail interpreting findings from statistical tests, regression models, and correlation analyses, providing insights into relationships between variables and informing conclusions or recommendations

Conclusion

In conclusion, AI offers promising avenues for enhancing education through personalized learning experiences. However, ethical considerations and collaboration between educators and developers are crucial for successful implementation and ongoing improvement.

2

4

6

1

3

5

Data Cleaning: Handling Missing Values

Data cleaning involves addressing missing values, ensuring data integrity and accuracy for effective analysis and decision-making.

Correlation Analysis

Correlation analysis includes Spearman and Pearson methods, assessing the strength and direction of relationships between variables, accommodating different data types and distributions.

Roadmap for Further Development and Deployment

The roadmap includes refining algorithms, optimizing efficiency, integrating real-time data, and rigorous testing before deploying the AI educational system.

DATA COLLECTION AND PREPROCESSING

DATA SOURCES

Public datasets like UCI Machine Learning Repository's Student Performance dataset (https://archive.ics.uci.edu/ml/datasets/Student +Performance)

DATA CLEANING

Remove any inconsistencies, missing values, and outliers from the raw data to ensure data quality.

FEATURE ENGINEERING

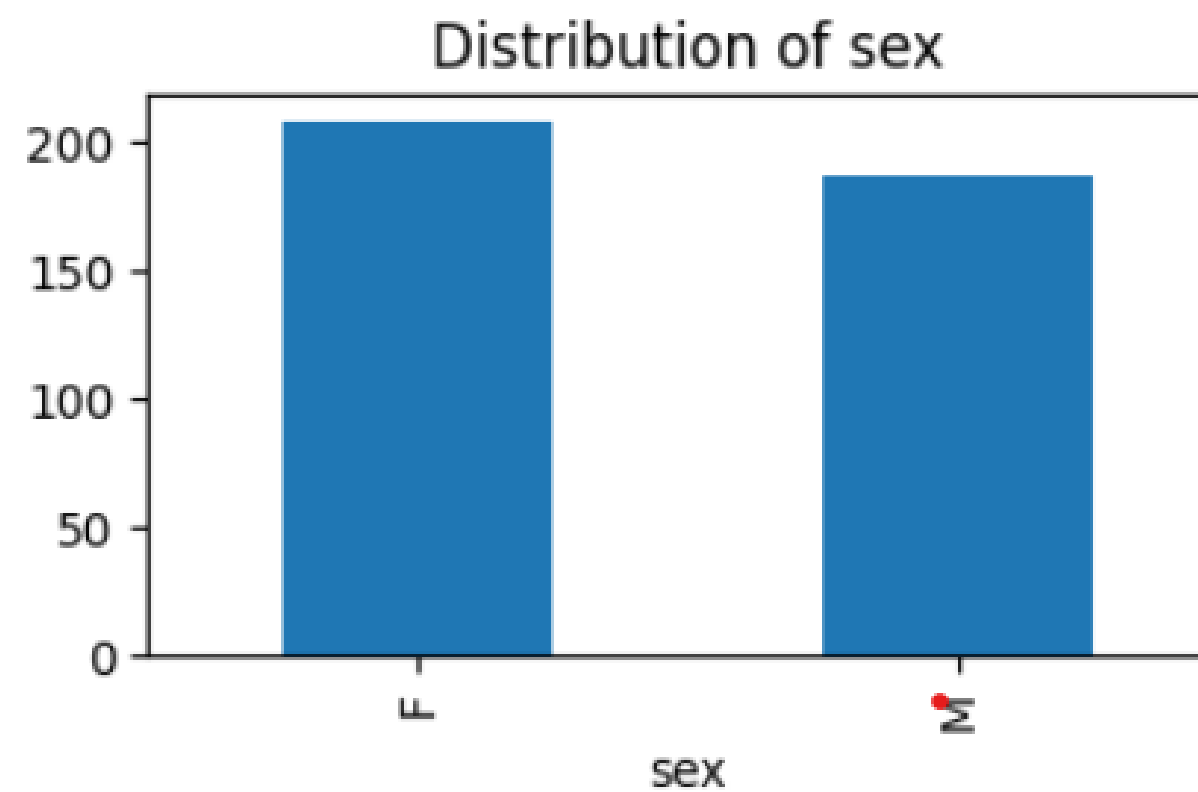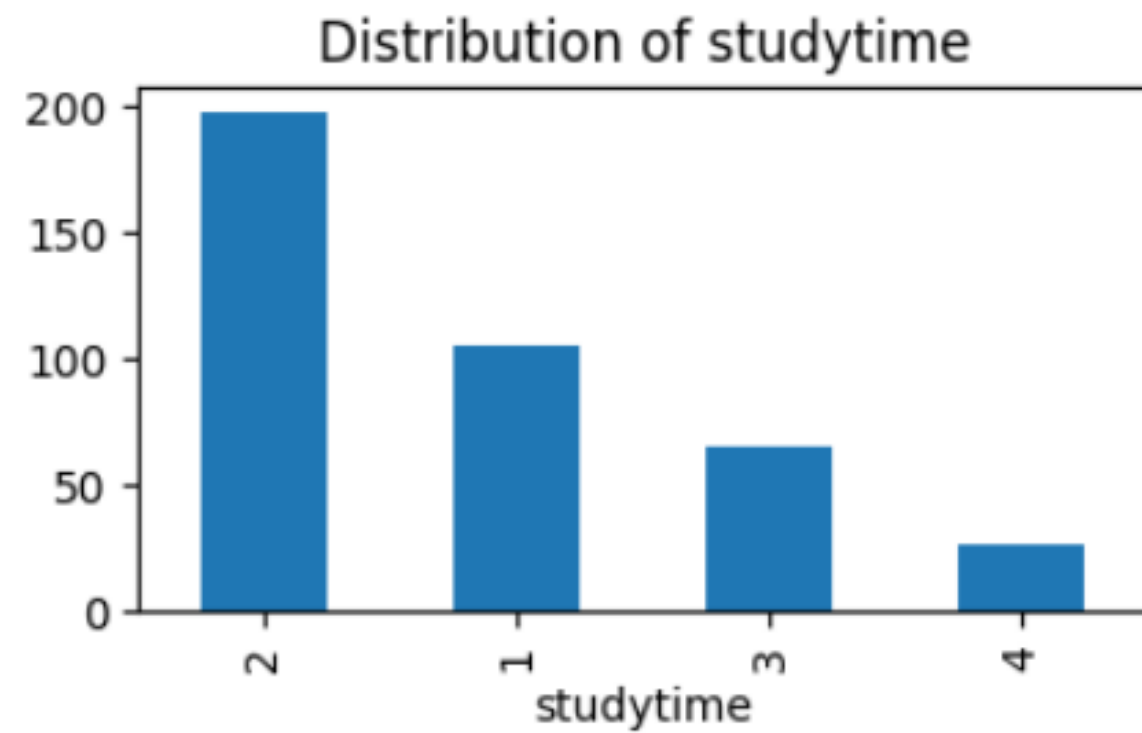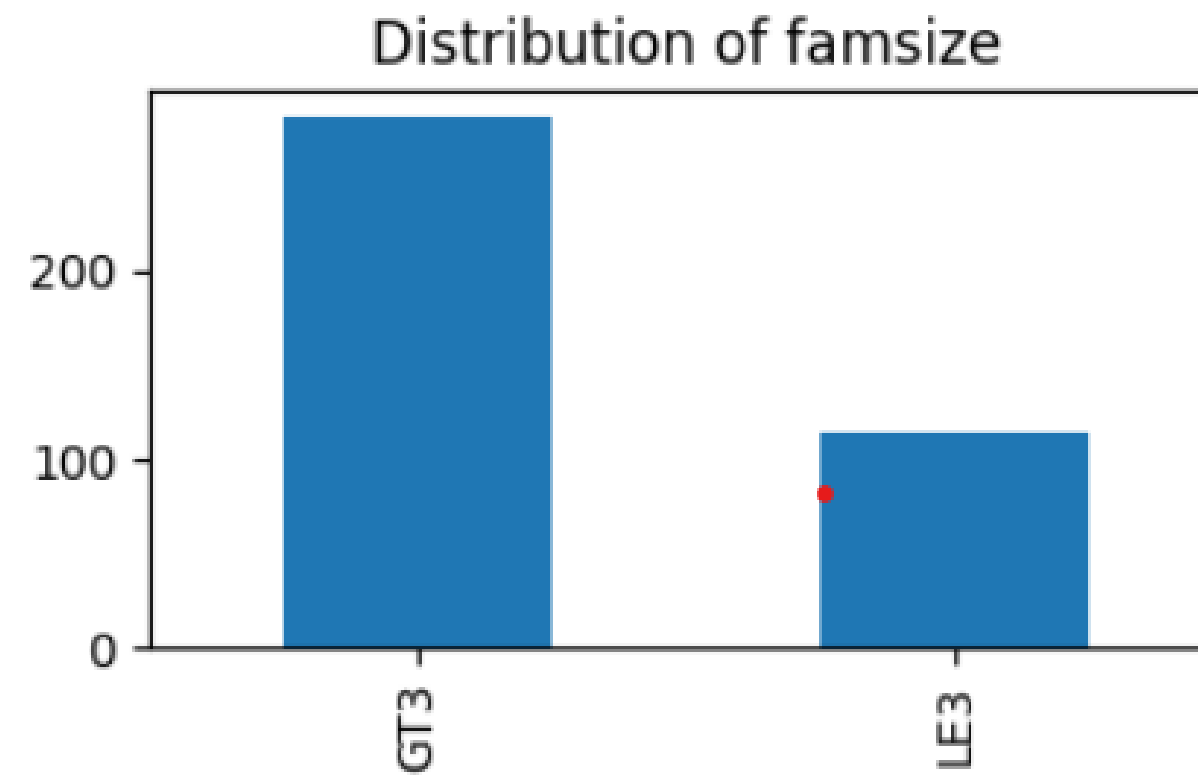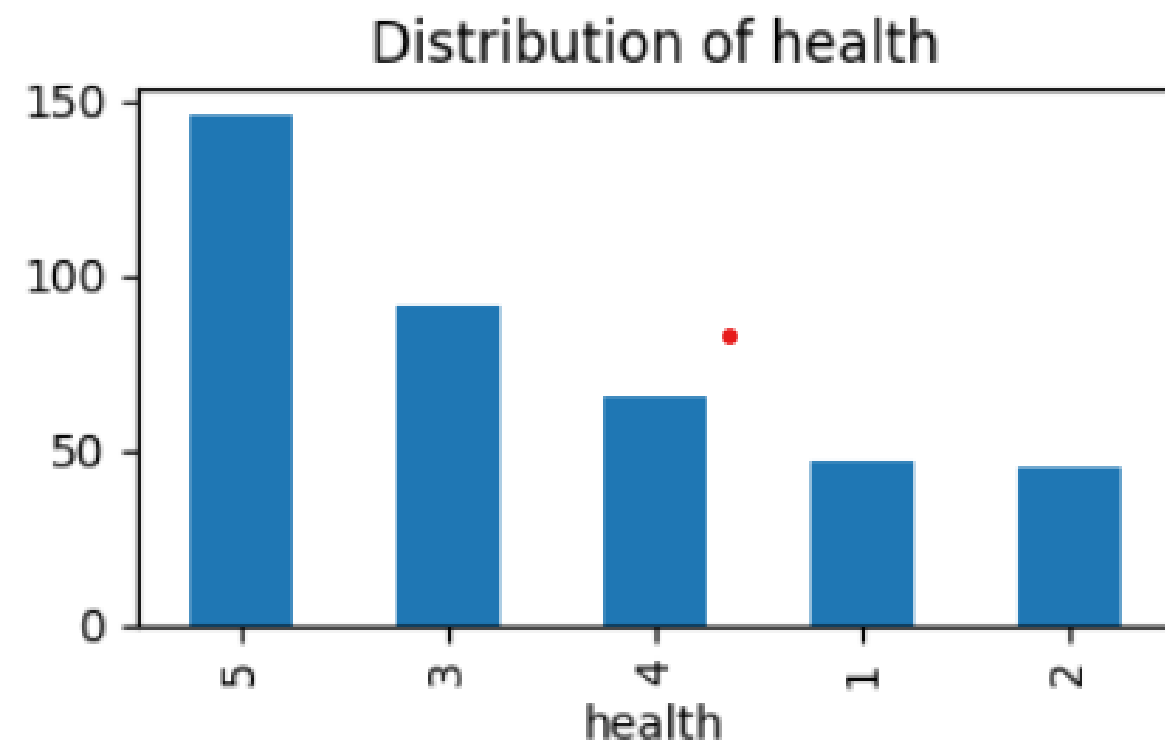Identify and extract relevant features from the data that could potentially impact student performance.

DATA
VISUALIZATION

Plots of the distribution of each feature using box plots, bar graph ,density plot and adjusts subplot spacing for better visualization.
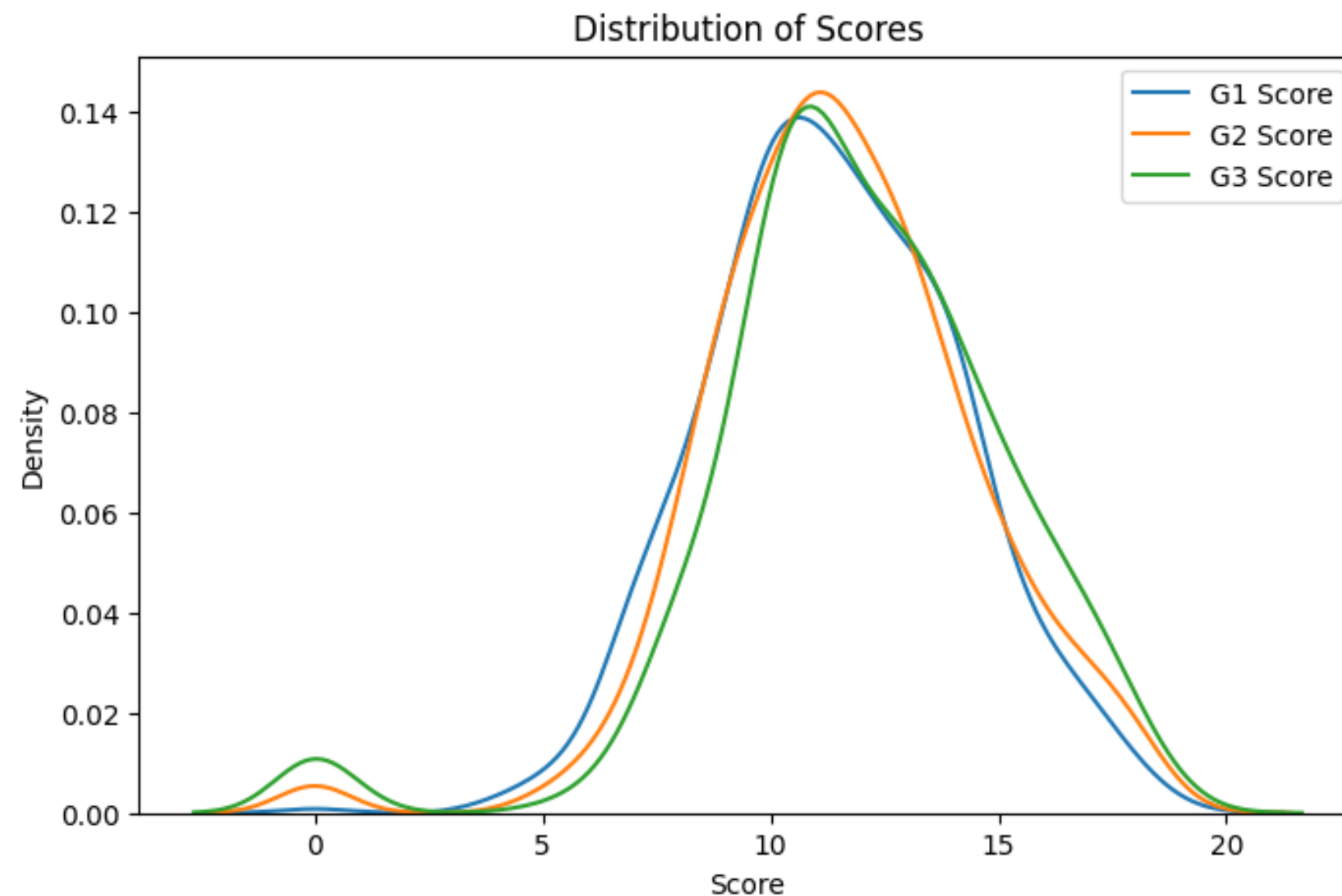
BAR GRAPHS

VISUALIZING STUDENT PERFORMANCE ACROSS ACADEMIC TERMS

This figure displays the distribution of student scores for three academic terms (G1, G2, G3) using kernel density estimation (KDE) plots. It offers a concise visualization of score distribution across the terms, aiding in understanding the spread and concentration of scores.

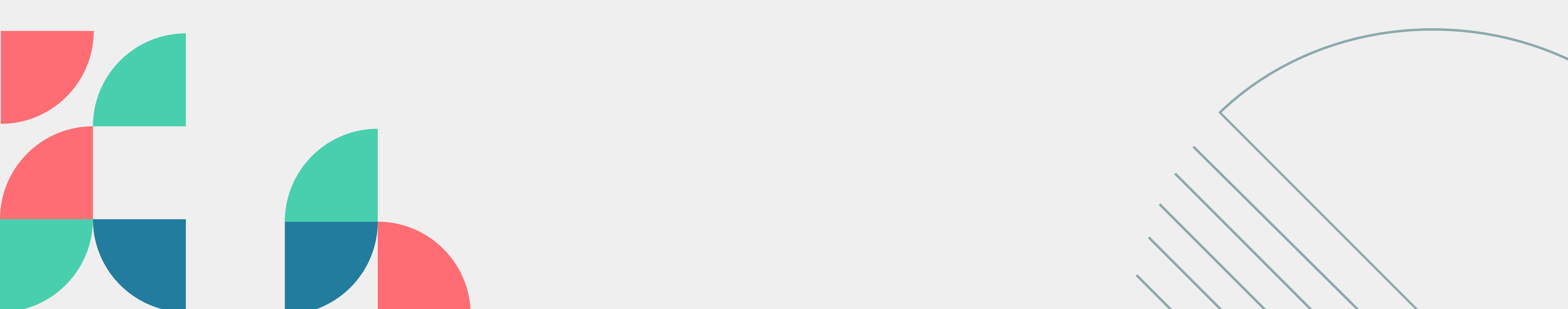# STUDENT ACADEMIC PERFORMANCE BY SEX: G1, G2, AND G3 SCORES

We have three box plots, one for each academic term (G1, G2, G3), showing student scores categorized by sex. It employs seaborn for plotting and matplotlib for subplot organization, ensuring a clear presentation of the data.
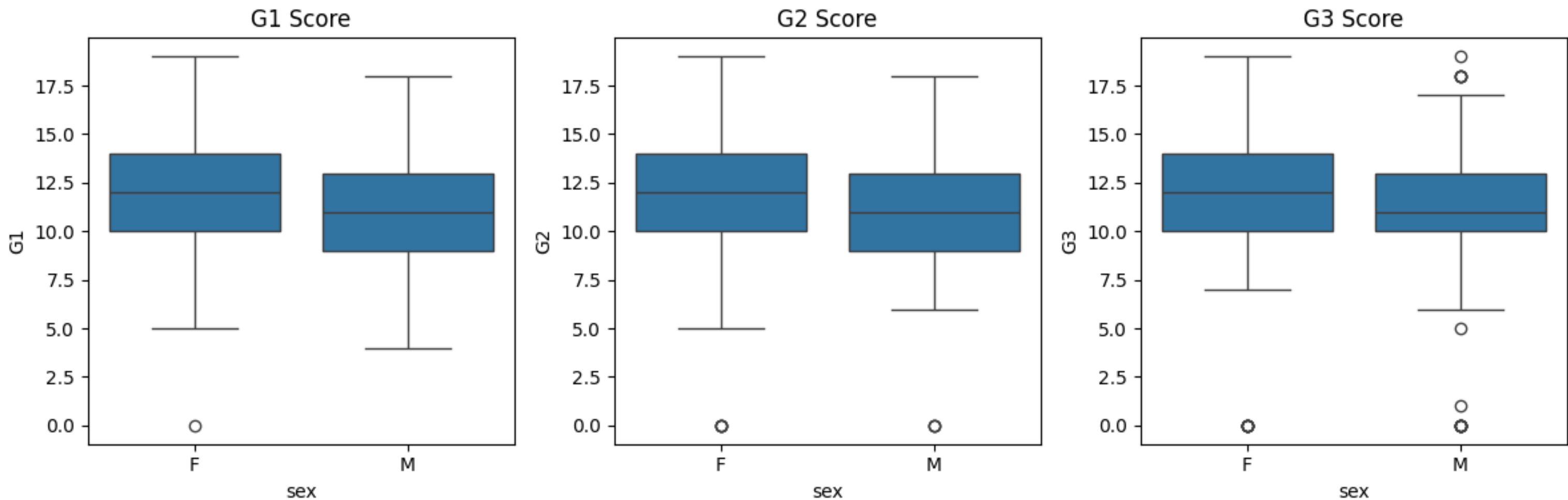
REGRESSION ANALYSIS

Regression analysis predicts numerical values based on input variables, revealing relationships between them and enabling predictions of how one variable changes with another.
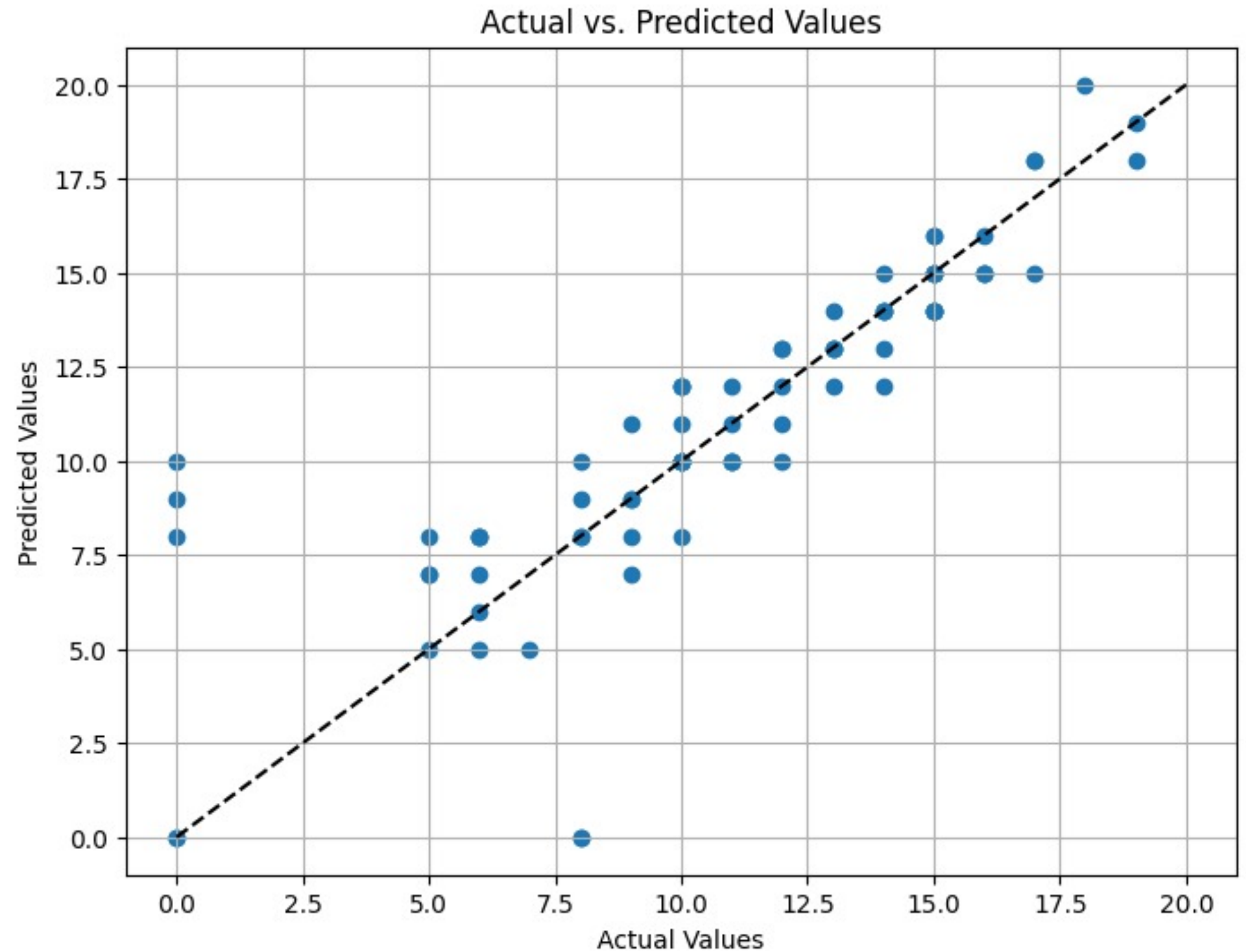
LINEAR REGRESSION MODEL

DATA PREPROCESSING

MODEL BUILDING

TRAINING AND EVALUATION

Actual vs. Predicted Values
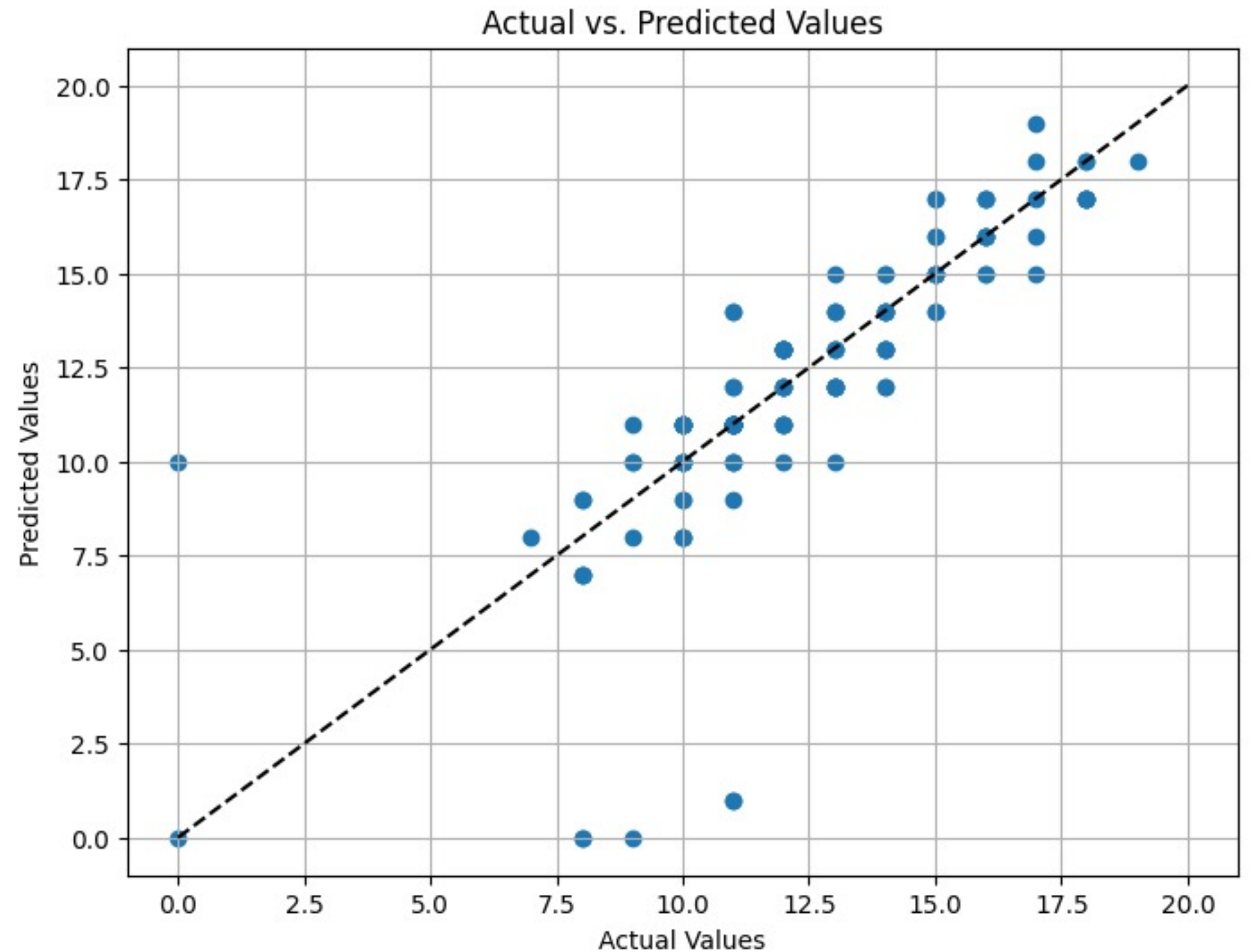
Portugese

DECISION TREE REGRESSION

DATA PREPROCESSING

MODEL BUILDING

TRAINING AND TESTING

MODEL EVALUATION

Mathematics

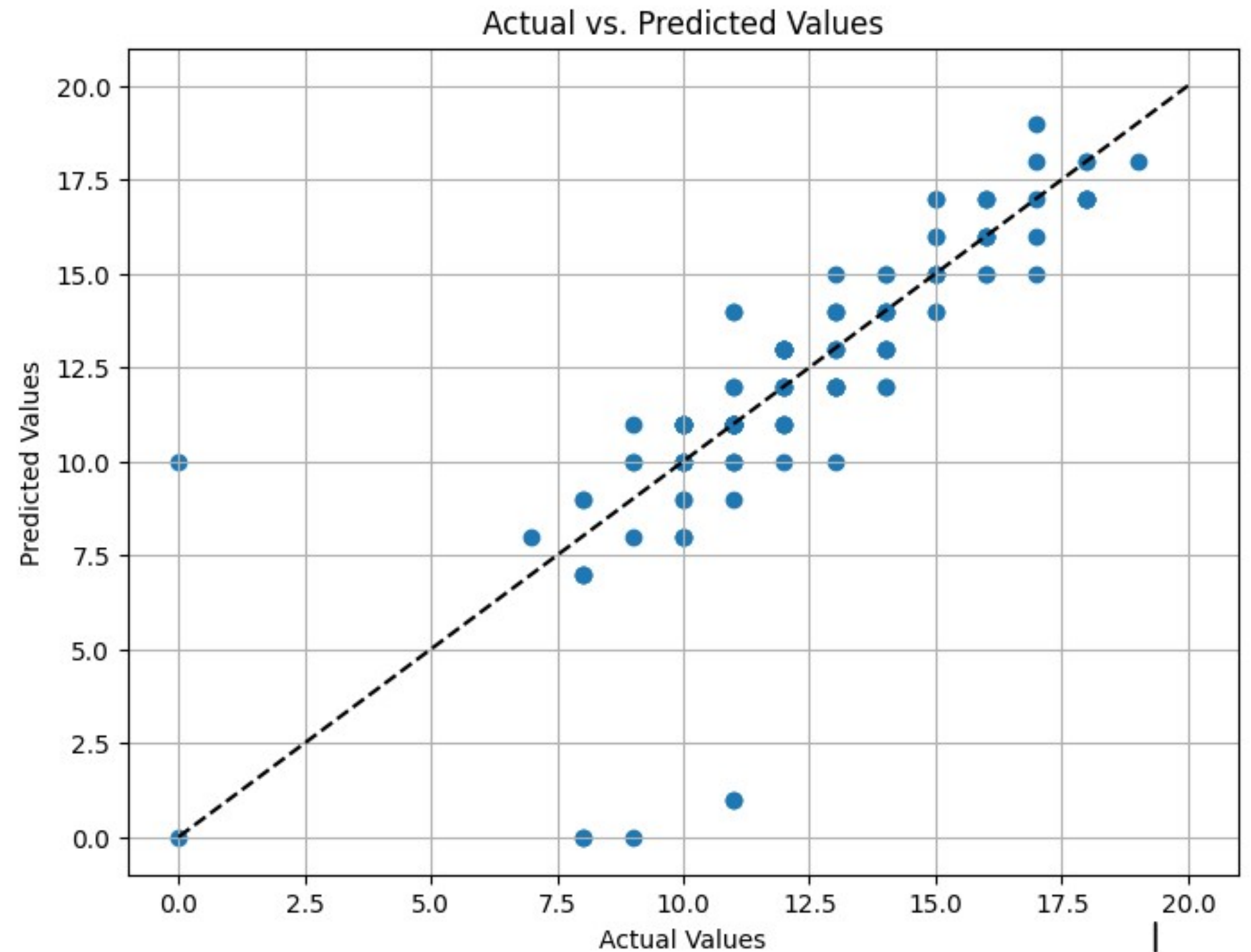# RANDOM FOREST REGRESSION

DATA PREPROCESSING

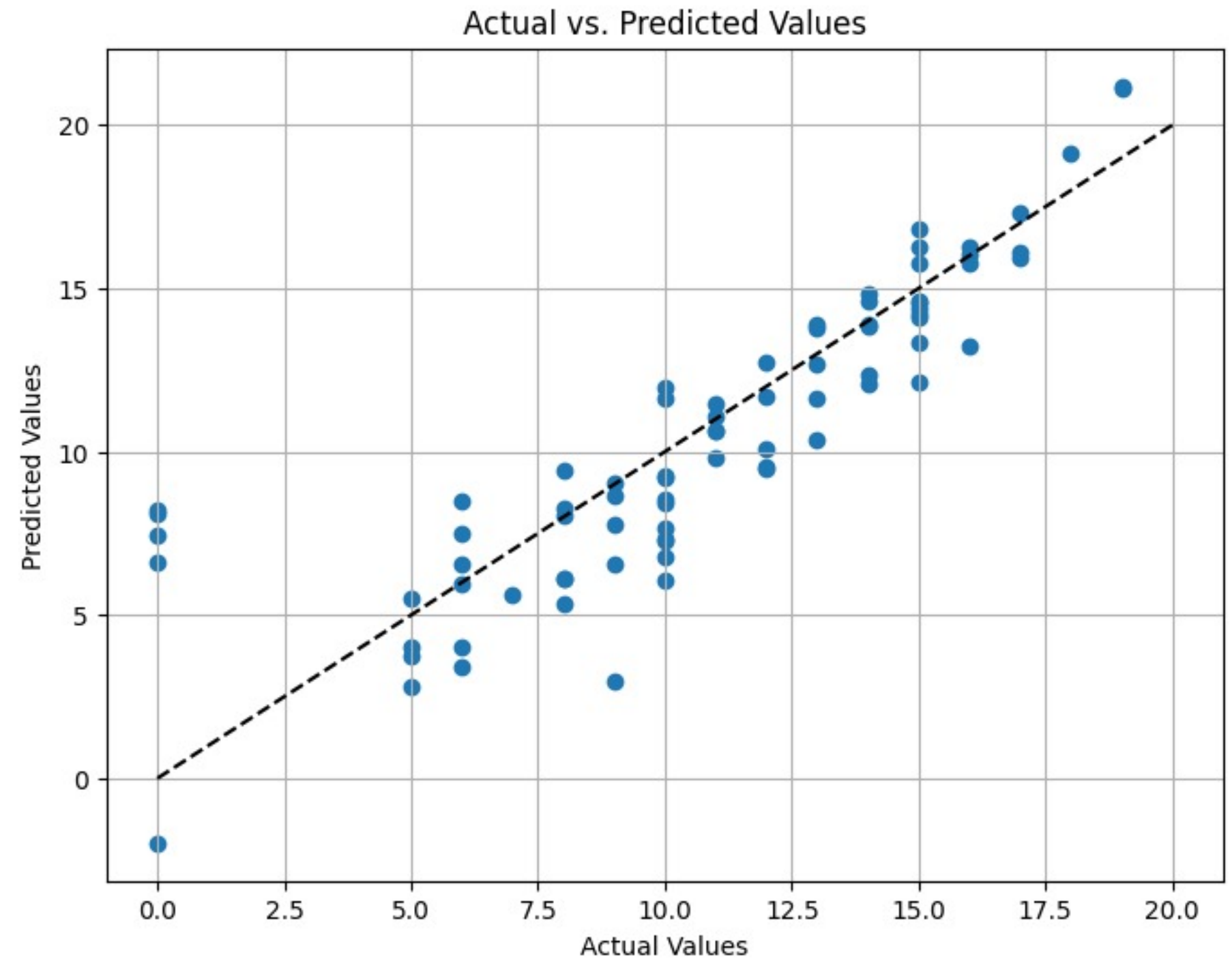**MODEL BUILDING**

TRAINING AND EVALUATION

Actual vs. Predicted Values

Portugese

SUPPORT VECTOR REGRESSION (SVR)

DATA PREPROCESSING

MODEL BUILDING

TRAINING AND EVALUATION

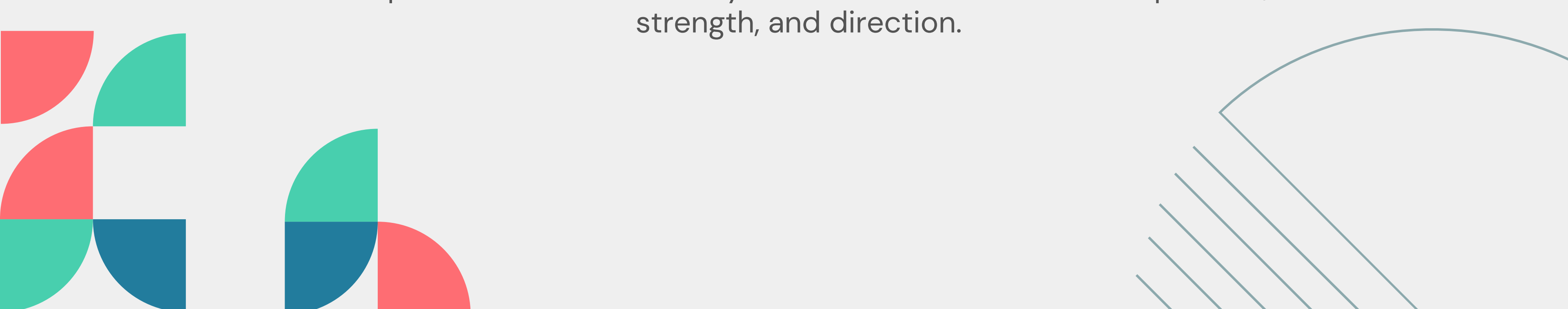Actual vs. Predicted Values

Mathematics

CORRELATION ANALYSIS

Correlation analysis evaluates the relationship between two variables, best suited for those showing a linear relationship. It's assessed visually through scatterplots. This bivariate analysis determines if a relationship exists, its strength, and direction.

SPEARMAN CORRELATION ANALYSIS OF STUDENT GRADES IN MATHEMATICS AND PORTUGUESE

Spearman correlation evaluates the strength and direction of the relationship between two variables, such as student grades in Math and Portuguese. It helps determine if there's a significant correlation between them, shedding light on how performance in one subject may affect the other.

Correlation Matrix Heatmap
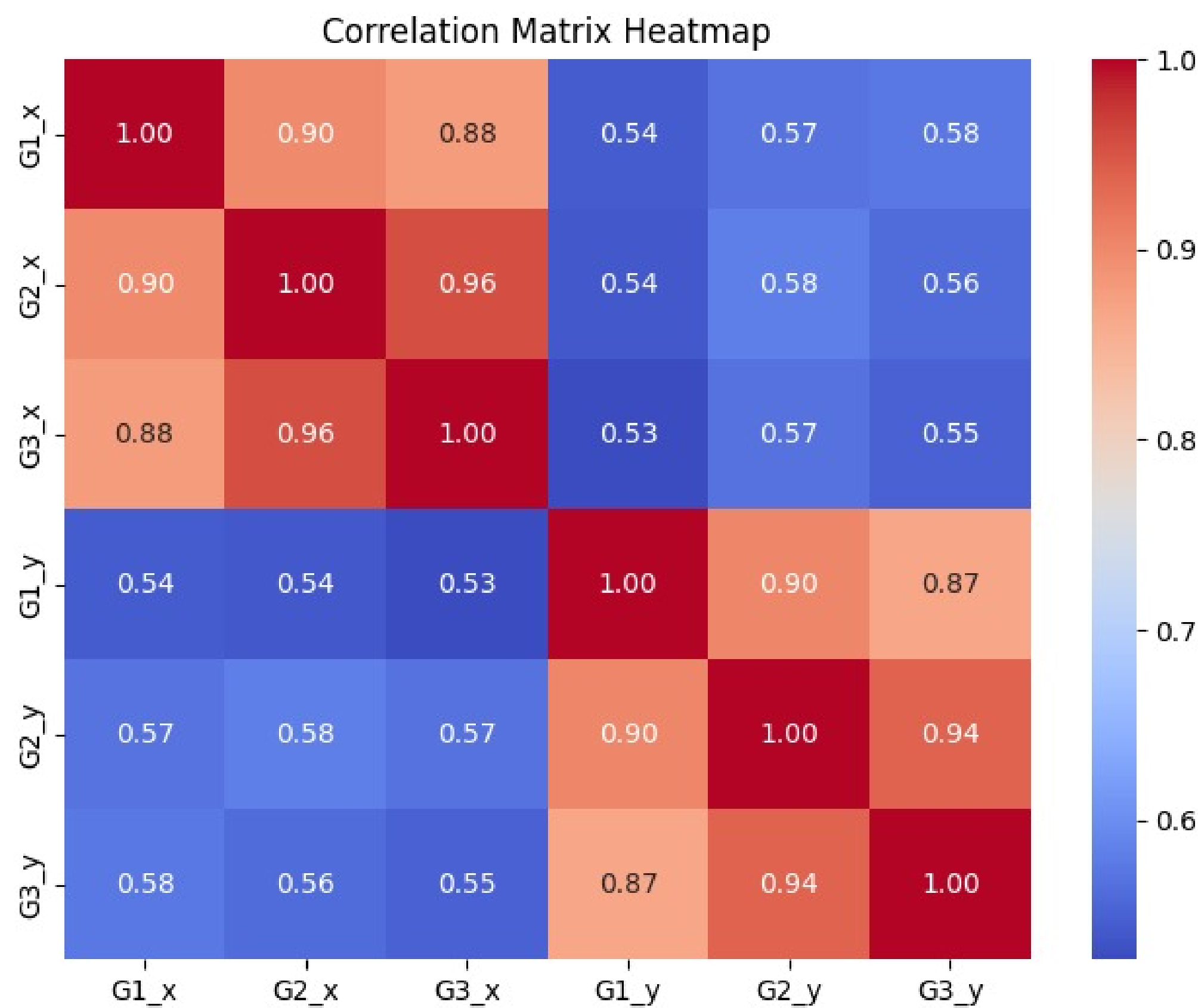
PEARSON CORRELATION ANALYSIS OF STUDENT GRADES IN MATHEMATICS AND PORTUGUESE

Pearson correlation assesses the strength and direction of the linear relationship between two continuous variables, such as student grades in Math and Portuguese. It's used to determine if there's a significant linear correlation between them in performance analysis.

Correlation Matrix Heatmap
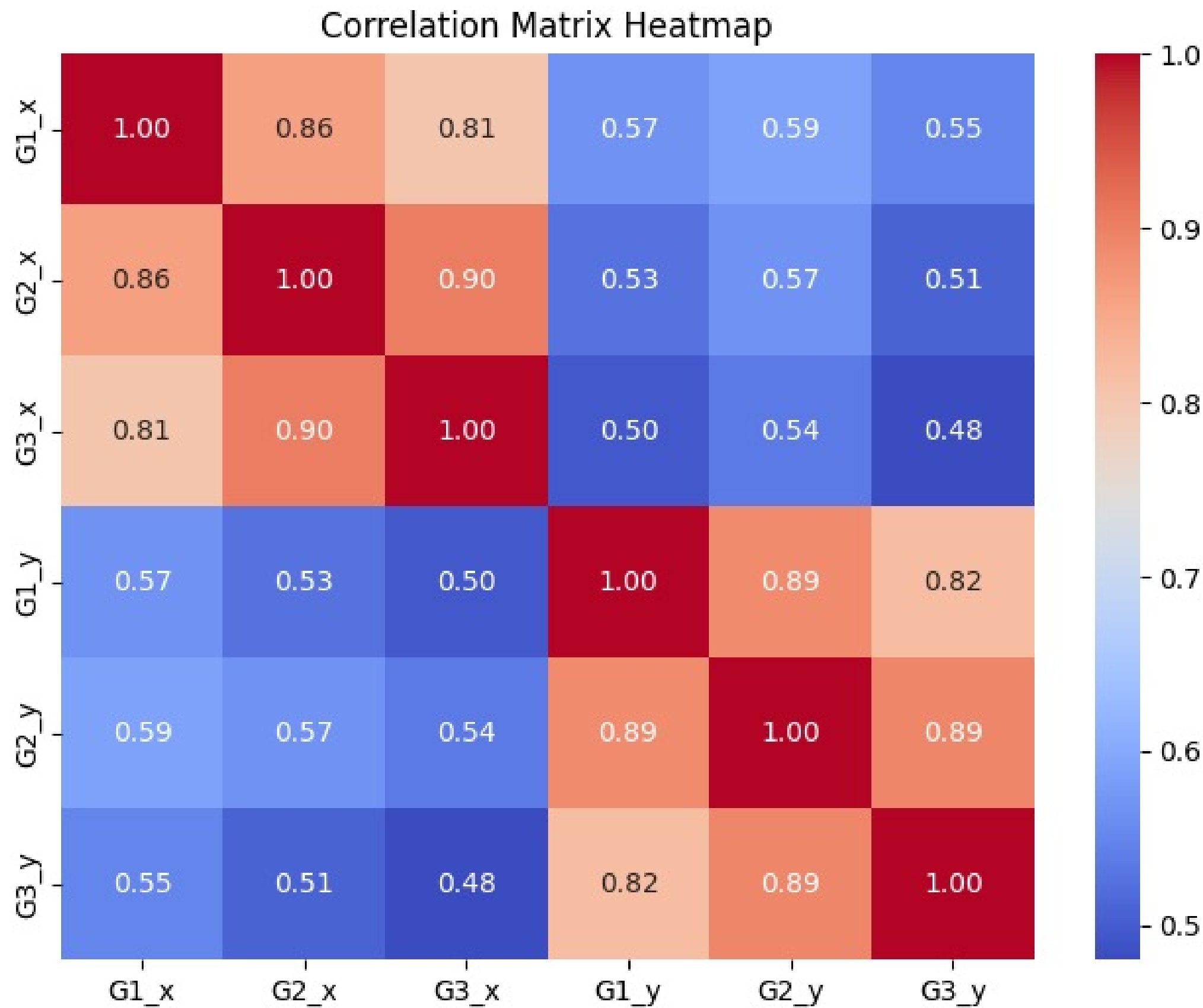
RESULTS & ANALYSIS

Among the regression models, the Random Forest Regression model demonstrates the highest prediction accuracy and robustness to nonlinear patterns in the data. Its ensemble nature allows it to capture complex relationships between input features and output grades, making it a suitable choice for predicting student performance.

- **Pearson Correlation:** Measures linear relationship between continuous variables, suitable for linear associations but may miss nonlinear ones.
- **Spearman Correlation:** Assesses monotonic relationship, robust to outliers and non-normal distributions, captures both linear and nonlinear associations.

ROADMAP FOR FURTHER DEVELOPMENT

1 **MODEL REFINEMENT AND OPTIMIZATION:**

identify areas for improvement. Explore advanced algorithms and techniques to enhance prediction accuracy and generalization.

2 **INTEGRATION OF REAL-TIME DATA STREAMS:**

Develop robust data pipelines to ingest, process, and integrate real-time data streams from educational platforms.

3 **CONTINUOUS MONITORING AND FEEDBACK MECHANISMS:**

Establish monitoring dashboards and performance metrics to track the deployed model's performance.

THANK YOU