# Statistics Assignment

In this assignment, you will perform a comprehensive statistical analysis of the Iris Species dataset. This dataset consists of measurements of four features of three different species of Iris flowers. Your task is to apply various statistical concepts to gain insights and make inferences about the data.

**Task 1:** Descriptive Statistics
Calculate and interpret the mean, median, and mode of each of the four attributes (sepal length, sepal width, petal length, and petal width) for the entire dataset. Calculate the standard deviation and variance for each attribute. What do these values tell us about the data?

**Task 2:** Correlation Analysis:
Calculate and interpret the correlation matrix for the four attributes by using a correlation table. Which attributes are positively or negatively correlated? How strong is the correlation?

**Task 3:** Hypothesis Testing
Perform a hypothesis test to determine if there is a significant difference in sepal length between the Iris setosa and Iris versicolor species. State your null and alternative hypotheses, conduct the test, and provide a conclusion

**Task 4:** Regression Analysis
Test the effect of sepal width on sepal length through a simple linear regression
Interpret the regression coefficients and provide the regression equation.

Bonus point: Create a scatter plot with the regression line overlay to visualize the relationship between sepal length and sepal width.

Submit a report that includes your findings, calculations, and interpretations for each task.