Data Cleaning, Analysis, and Visualization

Part A: Data Reviewing and Cleaning Exercises

- Check the structure and quality of the data provided. Provide a summary of your findings.

Part B: Summary Statistics

- Calculate mean, median, and standard deviation of University Exam Score by Sex.
- Which gender had a higher average?
- Calculate the correlation between Test 1, Test 3, and University Exam.
- Which pair of tests is most strongly correlated?

Part C: Data Visualization

- Visualize score distribution by Sex using any plot of your choosing
- Add appropriate labels and titles.
Histogram of Test 3 Scores
- Create a histogram showing distribution of Test 3 scores.
- Highlight any outliers.
Scatter Plot: Test 1 vs. University Exam
- Create a scatter plot to analyze relationship between Test 1 and University Exam.
- Color the points by Sex.

Part D: Model

Create a Clean Version of the Dataset
- After removing the identified issues from Part A
Build a Simple Linear Model
- Predict University Exam Score using Test 1 and Test 3 scores.
- What is the interpretation of the model coefficients?