Statistics Assignment (Total Marks: 20)

Objective:

To apply hypothesis testing, correlation, regression, and ANOVA techniques on real-world data and interpret the results with appropriate statistical justification.

Assignment Instructions:

1. **Select a Real-World Dataset (any one of the following):**
    - Public health (e.g., diabetes, heart disease, cancer survival)
    - Education outcomes (e.g., student performance, dropout rates)
    - Consumer behavior (e.g., online purchases, mobile phone usage)
    - Environmental data (e.g., air pollution, water quality, rainfall)
    - Any other area relevant to your thesis/research interest
2. **Apply at least THREE (2) from the following statistical techniques:**
    - **Hypothesis Testing** (1 mean, 2 means, 1/2 proportion, or 1/2 variances)
    - **Chi-Square Test** (Goodness of Fit or Contingency Table)
    - **Simple or Multiple Regression**
3. **Your report should include:**
    - **Introduction** (brief context of dataset and research question)
    - **Descriptive Statistics** (mean, variance, visualizations like histograms, boxplots)
    - **Methodology** (what tests and why they were chosen)
    - **Results and Interpretation** (statistical outputs, p-values, confidence intervals)
    - **Conclusion** (summary of findings in real-world terms)
4. **Use software:** R, Python, SPSS, or Excel to analyze the data. Include code/output screenshots where possible.

Evaluation Rubric (Total: 20 Marks):

Criteria	Marks
Relevance & Clarity of Research Question	2
Proper Application of Statistical Tests	6
Correct Interpretation of Results	4
Quality of Data Presentation (Tables, Graphs)	3
Use of Software and Reproducibility	2
Conclusion and Insights	3

Submission Deadline: [on exam day]