DATASET 6 – College Students' Academic Stress & Lifestyle

File: dataset_6.csv
Dependent variable (Y): StressScore
Independent variables (X): WeeklyScreentimeHours, CoCurricularHours, SleepQualityScore (+ Gender as subgroup)

Variable description

- ID – Student ID (1–50)
- Gender – 0 = Male, 1 = Female
- WeeklyScreentimeHours – Screen time hours per week
- CoCurricularHours – Hours per week spent on clubs/sports/other activities
- SleepQualityScore – Self-rated sleep quality (1 = very poor, 5 = very good)
- StressScore – Academic stress score (10–60, higher = more stress)

Suggested project questions

1. **Objective & Intro**: How are sreen time, co-curricular involvement, and sleep quality related to academic stress among college students?
2. **Descriptive statistics**
    A. Compute mean and SD of StressScore, WeeklyScreentimeHours, and SleepQualityScore.
    B. Compare mean StressScore between male and female students.
3. **Data visualization**
    A. Plot a histogram or boxplot of StressScore. Are there any outliers?
    B. Make a scatterplot of StressScore vs WeeklyScreentimeHours.
4. **Regression analysis**
    A. Fit a simple regression: StressScore (Y) vs WeeklyScreentimeHours (X). Interpret the slope: how much does stress increase per extra screen time in hour?
    B. Fit another model: StressScore vs SleepQualityScore. Is the slope negative (better sleep → less stress)?
5. **Subgroup comparison**
    A. Fit StressScore vs WeeklyScreentimeHours separately for males and females.
    B. Does the effect of screen time in hours on stress differ by gender?