Use the StudentsPerformance dataset and try the below questions:

Section A: Data Cleaning & Preprocessing

1. Check for missing values and handle them using:

    o Mean/median imputation for numerical.

    o Mode imputation for categorical.

2. Rename column names to remove special characters and use consistent casing (e.g., parental_education).

3. Generate descriptive statistics for numerical and categorical variables without using the describe function. (Write your own function).

4. Plot the distribution of math, reading, and writing scores. – Try out yourselves.

5. Create a new column average_score and categorize students into "Low", "Medium", "High" performers based on thresholds. – Hint: use pd.cut function

Additional: Visualization questions

6. Plot boxplots to compare scores by:

    o Gender

    o Test preparation course

7. Create a heatmap to visualize correlations between the scores.

8. Generate count plots (bar plots) for categorical features like gender, lunch type, and education level.