Homework #5

Objective: Gain hands-on experience with the apply family of functions (`apply`, `lapply,`, `sapply`, `vapply`, `tapply`, `mapply`) in R. Understand their use in different data manipulation and analysis scenarios.

Data: Use any publicly available dataset(s) relevant to environmental science, meteorology, or a field of your choice. The dataset should have a mix of numeric and categorical variables and be suitable for applying various statistical functions.

Tasks:

1- Data Exploration with apply: (20pts)

a. Explain the dataset you have chosen and the variables it contains.
b. Load your chosen dataset into R.
c. Use the `apply` function to calculate the mean and standard deviation for each numeric variable in the dataset. (**Bonus:** Normalize these numeric variables using `apply` **+10pts**)

2- List Processing with `lapply` and `sapply`: (20pts)

a. Create a list of several vectors, each representing different daily measurements of a particular pollutant.
b. Use `lapply` to calculate the median of each vector.
c. Then, use `sapply` to find the variance of each vector.

3- Factor Analysis with tapply: (20pts)

a. Group your data based on a categorical variable (e.g., city or date).
b. Use `tapply` to calculate the sum of a particular numeric variable within each group.

4- Safe Apply with vapply: (20pts)

a. Create a function that returns a numeric vector of fixed length.
b. Use `vapply` to apply this function to a list or vector, ensuring that the output has the correct length and type.

5- Parallel Processing with mapply: (20pts)

a. Create two or more vectors of equal length, each representing different environmental measurements.
b. Use `mapply` to calculate a combined metric (of your choice) for each set of corresponding elements from the vectors.

Bonus Challenge with rapply: (+20pts)

a. Create a nested list structure representing measurements from different locations and days.
b. Use `rapply` to apply a function at a specific level of the list (e.g., calculate a summary statistic for each day).

Assessment Criteria:

- Correctness and efficiency of the R code.
- Understanding of the apply functions and their appropriate usage.
- Clarity and thoroughness of the report.

Notes:

- Please, include your dataset in the same folder with your R script file or provide a link to the dataset.
- Write the methodology and rationale for using each *apply function* as *R comments* in the text file.