Assignment 3 Descriptive statistics

Due date: June 19, 11:59pm

Purpose:
Descriptive statistics provide basic understanding about a dataset. In this assignment, you will access data from an external file and calculate commonly used descriptive statistics for multiple variables in the dataset. You will also calculate descriptive statistics for sub-groups under the same variable.

Tasks: Please write an R script that performs the following operations in the order listed. At the beginning of each task, write a comment marking the task number. Name the file as: lab3-<your last name>.R

1. Make the following packages available: tidyverse, ggpubr, rstatix (You need to install these packages if they haven't been installed)
2. Download the 'happiness.csv' file from Blackboard. This dataset contains 'happiness score', 'health score', 'freedom score' and 'trust score' of countries from 4 regions: Africa, AsiaPacific, Europe, and PanAmerica
3. Import 'happiness.csv' dataset and save it to a data frame named 'happydata'. Check the structure of "happydata"
4. Browse and understand the data in each column
5. Central tendency: Calculate mean, median, mode for Happiness.Score
6. Spread: Calculate standard deviation, variance, min, max, range, and quantile for Happiness.Score
7. Use the sapply() function to calculate the mean and standard deviation for all 4 scores.
8. Use the summary() function to get a summary statistics for the entire dataset
9. Calculate the mean and standard deviation of Happiness.Score for different regions. What do you learn from the result?

Submission:
Submit your assignment by uploading the actual R script itself to ELMS via the Lab3 submission link.