STAT 217: Homework 1

Due Friday, Aug 28th beginning of class

Background: A study of the perceived stress levels of prisoners was conducted that involved measuring the Perceived Stress score that ranges from 0 to 56 points with higher scores related to more perceived stress. The prisoners were measured at the beginning and end of the study, with some of the prisoners involved in an exercise program and others not. We are interested in the changes in perceived stress levels over the course of 22 weeks of imprisonment for offenders. Specifically, we will calculate a Difference in stress (after 22 weeks – before 22 weeks) to generate a variable named Difference in the stress data frame.

Your task is to summarize those results using the provided code. In order to get started with R using R-studio, you will need to open R-studio and follow the instructions below. Some of the code is given to you, and some you will have to write yourself. The goal is to produce numerical summaries of the Differences, a histogram of the Differences, and a boxplot of the Differences.

Steps:

- Open a new R-script in RStudio. Go to File - New File - R Script. Save the script to your computer. Call it something meaningful like "`Homework 1`"
- Import the data into RStudio. The data file, `stress.csv`, is posted on D2L under DataSets. Save this file in an appropriate folder on your computer and follow the instructions on page 11 of the textbook.
- Notice a new tab with a data spreadsheet opens after import. Exit out of this tab and return to your "`Homework 1`" R script.

1. Your writeup starts here. In your writeup, print the code that you run and the output that you get for each of the following. Write the code in boldface font and the output in normal font. Follow the example I show in part (a).

   (a) At this point, you should have a new data.frame loaded into R called stress. Let's look at the data. Type `stress` into your R script. Put your cursor on this line and press Ctrl-Enter to run the code. You should see the dataset pop up in your console.

   Example Writeup
   Code: `stress`
   Output:

```
 Subject Group        PSSbefore       PSSafter
1   A1    Sport           25         13
2   A3    Sport           17         15
3   A4    Sport           12         13
4   A5    Sport           21         22
5   A6    Sport           29         25
6   A7    Sport           28         24
```

7	A8	Sport	21	19
8	B1	Sport	18	19
9	B2	Sport	20	24
10	B3	Sport	29	21
11	B4	Sport	27	22
12	B5	Sport	44	33
13	B6	Sport	23	8
14	B7	Sport	18	12
15	B8	Sport	27	30
16	A2	Control	0	16
17	C1	Control	30	27
18	C2	Control	12	31
19	C3	Control	29	21
20	C4	Control	25	33
21	C6	Control	2	9
22	C7	Control	6	26
23	C8	Control	20	20
24	C9	Control	27	28
25	C10	Control	14	21
26	C11	Control	15	29

(b) Now, follow the instructions of page 12 of the textbook to lget R to print the first six lines of the stress dataset. Print your code and output in your writeup.

Head(stress)

Subject	Group	PSSbefore	PSSafter
1	A1 Sport	25	13
2	A3 Sport	17	15
3	A4 Sport	12	13
4	A5 Sport	21	22
5	A6 Sport	29	25
6	A7 Sport	28	24

(c) Now, follow the instructions to get R to print the last six lines of your dataset. Make sure to print your code and output in your writeup.

Tail(stress)

Subject	Group	PSSbefore	PSSafter
21	C6 Control	2	9
22	C7 Control	6	26
23	C8 Control	20	20
24	C9 Control	27	28
25	C10 Control	14	21
26	C11 Control	15	29

(d) Recall that we are interested in the difference in stress levels (after-before) for each subject. The following line of code adds a new column to your dataset called Differences. Run this line of code and then look at the stress dataset again. Print the new stress dataset in your writeup.

```
stress$Differences<-stress$PSSafter-stress$PSSbefore
```

Differences (-12, -2, 1, 1, -4, -4, -2, 1, 4, -8, -5, -11, -15, -6, 3, 16, -3, 19, -8, 8, 7, 20, 0, 1, 7L, 14)

(e) For i-v, follow the instructions on page 14 of the text.

i. Print only the Differences column

```
 [1] -12  -2   1   1  -4  -4  -2   1   4  -8  -5 -11 -15  -6   3
[16]  16  -3  19  -8   8   7  20   0   1   7  14
```

ii. Print the mean and standard deviation of the differences.

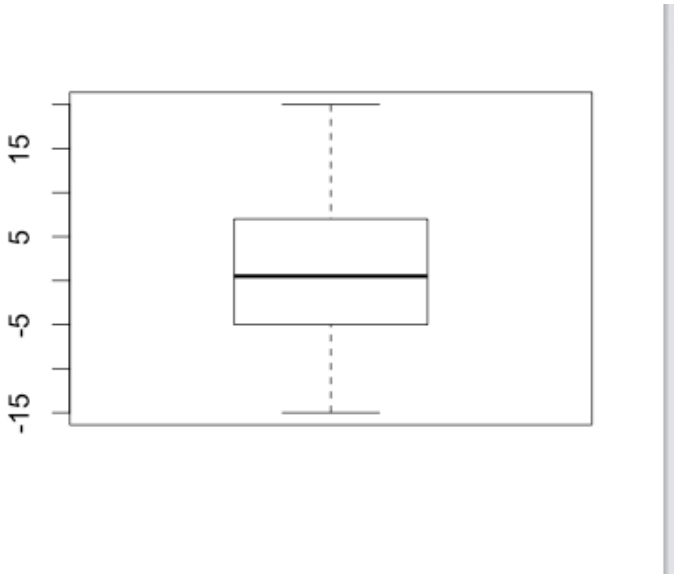
Mean: [1] 0.8461538 SD: [1] 9.194313

iii. Print the five number summary of the differences. The book says to use the function `favstats`, but instead use the function `summary`.

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
-15.0000  -4.7500   0.5000   0.8462   6.2500  20.0000
```

iv. Make a histogram of the differences

v. Make a boxplot of the differences.



2. In a short written report, summarize the results, focusing on describing the shape of the distribution of the differences and whether you think the perceived stress levels changed over the course of the study (was the observed change very different from 0?). Refer to the figures and summary measures you found in part (e).

   The mean difference in perceived stress in prisoners after 22 weeks was 0.846 of a point on a scale of 56. The standard deviation was 9.19. The histogram displaying the differences was slightly skewed right. The observed change was not very different from 0, though there were some extreme changes in some prisoners. For example, the greatest reduce in stress for one prisoner was 15 points. The greatest increase in stress for another prisoner was 20 points. Most of the prisoners showed slight increases or decreases in stress.