STAR 511 HW #1

See Canvas Calendar for due date.
12 points total

Important reminders:

- Use RStudio.
- Use the R markdown HW template which is available from Canvas.
  - Add or delete sections and code chunks as needed.
  - With the HW template all R code is shown in an appendix at the end of the assignment.
- Download the datasets from the Ott & Longnecker companion site (see Canvas > Modules > Week 1 or Additional Resources). Download the zip file and then **unzip the file**. We will use the CSV files ("**ASCII-comma**"). The file extension is .TXT even though the files are actually CSV!
- **For ease of importing, I recommend copying the needed data files into the same folder as your HW markdown document**. If this is not the case, you will need to specify the full file path location.
- Below we use read.csv(, `quote = "'"`) to import the data. **The quote option is used because the column names in the original data are (single) quoted.**
- Always look at the data after importing! This can be done within RStudio or using str(). **Check the exact column names and modify below code as needed.**
- **S**ubmit HW through Canvas in doc, docx or pdf format. Students generally have the easiest time knitting to a Word document.
- See RHelp document for additional details and suggestions.
- If you have questions, consider attending in person or Zoom meeting for individual help so we can screen share.

1. Use the data described in Problem 3.34 below regarding resting pulse rates:

> **3.34** The following data are the resting pulse rates for 30 randomly selected individuals who were participants at a 10K race.
>
> 49 40 59 56 55 70 49 59 55 49 58 54 55 72 51
> 54 56 55 65 57 61 41 52 60 49 57 46 55 63 55

From the files you downloaded above, you will find the data under CH03, named ex3-34.txt. Read this file into RStudio, graph and summarize the data. Here is code that will read the data:

```
pulse <- read.csv("ex3-34.txt", quote = "'")
```

A. Use the str() function to display the data structure. Write a sentence or two describing what the output tells you.
B. Construct a histogram of pulse rate.
C. Give the mean and median for pulse rate.

2. Use the data described in Problem 3.7 below regarding survival times for two therapies:

3.7 The survival times (in months) for two treatments for patients with severe chronic left-ventricular heart failure are given in the following tables.

Standard Therapy							New Therapy						
4	15	24	10	1	27	31	5	20	29	15	7	32	36
14	2	16	32	7	13	36	17	15	19	35	10	16	39
29	6	12	18	14	15	18	27	14	10	16	12	13	16
6	13	21	20	8	3	24	9	18	33	30	29	31	27

You will find the data under CH03, named ex3-7.txt. Notice that the two therapies are in two different columns. (There are ways to reformat, but we will work with the data "as is" for now.)

A. Use the str() function to display the data structure. Write a sentence or two describing what the output tells you.
B. Construct side-by-side boxplots showing the survival times for each therapy.
C. Give the mean and standard deviation for <u>each</u> of the therapies.