Statistics 147 Assignment #3
Summer 2020; 35 pts

DUE DATE: Monday, July 13, 2020, by 10:00 pm

GENERAL INSTRUCTIONS:

♣ **No late papers will be accepted for credit!**

♣ Your write-up should be neat, well-organized and concise but complete. Your write-up should include a cover page which includes your name, the last 4 digits of your student ID. Your write-up must be typed. Use a word processor such as Word or LATEX, etc. No hand-written papers will be accepted for credit!

♣ Be sure to include copies of any/all charts, graphs and output.

♣ You will need the following data file, that is available for download on **Ilearn** under **Data Files**

agility.dat

♣ For your SAS session, name your program file *hwk3_su20.sas*.

(i) **(2 pts)** Include the following titles:

```
title1 'Statistics 147 Assignment #3';
title2 'Summer 2020';
title3 'Your Name';
title4 'SAS Question XX'    where XX = question number
title5 'SubPart (xx)';     /* where xx is something like ii */
```

(ii) **(4 pts)** Be sure to include documentation in your SAS program.

(iii) **(2 pts)** Be sure to turn in your SAS program and output!

♣ For your R session,

♠ **(2 pts)** be sure include the following titles (as comments)

```
# Statistics 147 Assignment #3
# Summer 2020
# Your Name
# R Question ZZ (where ZZ = question number)
# Subpart XX (where XX = subpart number, something like ii)
```

♠ **(2 pts)** Be sure to include a copy of your R code. (You may copy and paste it from the R Console into Word, LATEX, etc.)

♠ **(3 pts)** Be sure to include documentation in your R code.

The Questions

R

1. **(10 pts total)** (Use **R** for this problem!) Linda has three dogs (Cody, Dusty and Shadow) that she is training for a national agility championship. Linda records their times to finish the course (in seconds) for 30 runs. Lauren does not believe there is a significant difference in mean finishing times between the three dogs. To test this claim, Lauren takes three independent random samples of 8 times for each of the three dogs, yielding the following data:

```
File name: agility.dat
Cody Dusty Shadow
75      69      83
85      79      93
70      77      87
79      51      72
73      53      82
81      69      72
84      59      62
70      64      77
```

NOTE:

- ♠ The data is located in a datafile named **agility.dat**.
- ♠ The headings are included in the data file. The actual data begins on line 3.
- ♠ **Assume** the *Cody* is dog 1, *Dusty* is dog 2, and *Shadow* is dog 3.
- ♠ μ_i = true average finishing time for dog i and σ_i^2 = variance of the finishing time for dog i, $i = 1, 2, 3$

(i) Write the R code to read in and print out the data. (Done as part of Assignment #2.)

(ii) Add the appropriate lines of code to make the columns accessible individually and obtain the column headers. (Done as part of Assignment #2.)

(iii) [**NEW**] Using **R** to generate the appropriate output, test whether the true mean finishing time for **Cody** is more than 75 minutes. Use $\alpha = 0.05$. (5 pts)

(iv) [**NEW**] Using **R** to generate the output, find and interpret a 96% confidence interval for the true mean finishing time of **Cody**. (5 pts)

SAS

1. **(10 pts total)** Refer to **R** Question 1.

(i) Write the SAS code to read in and print out the data. Use **agility1** as your temporary SAS dataset name. Use nested do loops! (**NOTE: DO NOT COPY AND PASTE THE DATA INTO YOUR SAS PROGRAM. READ THE DATA IN FROM THE EXTERNAL DATA FILE!**) (Done as part of Assignment #2.)

(ii) Using if-then-else structures, name the dogs as follows: (Done as part of Assignment #2.)

dog	1	2	3
name	Cody	Dusty	Shadow

(iii) Add the appropriate lines of code to sort the data by the **name** of the dog. (Done as part of Assignment #2.)

(iv) Add the appropriate lines of code to your program to generate the mean, standard deviation and median for each of the dogs. (Done as part of Assignment #2.)

(v) Create a new temporary SAS data set, named **onlyC**, in which the data is restricted to the times of *Cody*. (Be sure to print the data as a check.) (Done as part of Assignment #2.)

(a) [**NEW**] Test whether the true mean time of the **Cody** team is 75 minutes. Use $\alpha = 0.05$. (5 pts)

(b) [**NEW**] Find and interpret a 98% confidence interval for the true mean finishing time for the *Cody* team. (5 pts)