Statistics 147 Assignment #5

Summer 2020; 35 pts

DUE DATE: Monday, July 20, 2020 by 5:45 pm.

General Instructions:

- Templates for the solution (in Word and LaTeX) to this assignment has been uploaded to Blackboard. They were designed to minimize your typing of the solution. You may use them if you like. They are located under **Course Materials → Handouts → Templates**. (Be sure if you do not use one of these templates that you include all the steps in the tests of hypothesis!)
- 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!
- No late papers will be accepted for credit.
- For your **SAS** session, name your program file **hwk5_su20.sas**. (Modify your SAS program from Assignment #4 to complete this assignment!)

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

```
title1 'Statistics 147 Assignment #5';
title2 'Summer 2020;
title3 'Your Name';
title4 'SAS Question XX';  /* where XX = question number  */
title5 'SAS Question Subpart YY';   /*  where YY = subpart, like (i)  */
```

  ♠ **(2 pts)** Be sure to turn in your SAS program and output! You may include the SAS code throughout the assignment, but also include the entire SAS code at the end of the assignment.

  ♠ **(2 pts)** Be sure to include comments in your SAS program.

- For your **R** session,

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

```
5
> # Summer2020
> # Your Name
> # R Question XX (where XX = question number)
> # R Question Subpart YY (where YY = subpart, like (i))
```

  ♠ Type and save your code in an **R Script** named **asmt5_RScript_su20_XX**, where **XX** = initials of your name. Attach a copy of your **R Script** to the end of your assignment. (This assignment is an extension of Assignment #4. Open your **R Script** from Assignment #4 and make modifications/additions.)

  ♠ **(2 pts)** Be sure to include a copy of your R code from the **R Console** (so that the code and results are printed). (You may copy and paste it from the R Console into Word, Notepad, LaTeX, etc.)

- **DATA Files:** You will need the following data file: **agility.dat**. (You should have downloaded this file for Assignments #3 and #4. It is located on Blackboard under **Data Files**.)

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, $\quad i = 1, 2, 3$

1. **Using R, complete the following:**

   (i) Test for underlying normality for each of the dogs. (Use Anderson-Darling Test.) (4 pts)

   (ii) Test for equality (homogeneity) of variances. (Use Bartlett's test.) Use $\alpha = 0.05$. (2 pts)

   (iii) Perform the appropriate test(s) of hypothesis to determine whether one can conclude that at least one of the dogs has a significantly different mean finishing time. Use $\alpha = 0.05$. (4 pts)

   (iv) If appropriate, use **Tukey's test** *and* the **p-value method** to determine which mean(s) is(are) significantly different. (Be sure to justify your answer!) (3 pts)

2. **Using SAS, complete the following.** Modify your existing **SAS** program file (from Assignments #3 and #4) to complete the following.

   (i) Test for underlying normality for each of the dogs. (Use Shapiro-Wilk Test.) (4 pts)

   (ii) Test for equality (homogeneity) of variances. (Use Bartlett's test.) Use $\alpha = 0.05$. (2 pts)

   (iii) Perform the appropriate test(s) of hypothesis to determine whether one can conclude that at least one of the dogs has a significantly different mean finishing time. Use $\alpha = 0.05$. (4 pts)

   (iv) If appropriate, use **Tukey's test** *and* both the **grouping** *and* **confidence interval** methods to determine which mean(s) is(are) significantly different. (Be sure to justify your answer!) (3 pts)