The due date is on Canvas.

Be sure to name your assignment in the following format:

Lastname_Firstname_S_HW_6.pdf

Only .pdf files will be accepted. You should be capable of compiling directly to a PDF file and not to any other file type beforehand.

Use complete sentences and show a "reasonable" amount of algebraic work. (Remember I grade the entire solution not just the answers.) Final answers should include the correct units.

Remember to use complete sentences/equations/etc, and context should be clear (answer the word problem).

Your work should be rather professional. Do the problems in order. You will be docked 5 points for the overall assignment if the homework isn't properly organized.

Any plots you make should be large enough so that the important aspects can be clearly seen. Every plot should include the appropriate x- and y- labels along with other important labels such as the scale or group identifier. Every plot should include a title.

If asked about the probability of a particular event, you should write

$$P(\text{event name or label}) = \text{appropriate amount of work} = \text{answer}$$

Unrelated problems are separated by a horizontal line.

By participating in this class, you must be aware of the course and university policies with regards to the VT Honor Code.

In particular, you acknowledge and support the statement:

"I have neither given nor received unauthorized assistance on this assignment. The work I am presenting is ultimately my own. "

---

1. In a 1994 Senate subcommittee hearing, an executive of a major tobacco company testified that the accusation that nicotine was added to cigarettes was false. Tobacco company scientists stated that the amount of nicotine in cigarettes was completely determined by the size of the tobacco leaf, with smaller leaves having greater nicotine content. Thus, the variation in nicotine content in cigarettes occurred due to a variation in the size of the tobacco leaves and was not due to any additives placed in the cigarettes by the company. Furthermore, the company argued that the size of the leaves varied depending on the weather conditions during the growing season, over which they had no control. To study whether smaller tobacco leaves had a higher nicotine content, a consumer health organization conducted the following experiment. The major factors controlling leaf size are the temperature and the amount of water received by the plants during the growing season. The experimenters created four types of growing conditions for tobacco plants. Condition A was average temperature and rainfall amounts. Condition B was lower than average temperature and rainfall conditions. Condition C was higher than average temperature with lower than average rainfall. Finally, condition D was higher than average temperature and rainfall. The scientists then planted 10 tobacco plants under each of the four conditions in a greenhouse where temperature and amount of moisture were carefully controlled. After growing the plants, the scientists recorded the leaf size and nicotine content, which are given here:

Plant	A Leaf Size	B Leaf Size	C Leaf Size	D Leaf Size
1	27.7619	4.2460	15.5070	33.0101
2	27.8523	14.1577	5.0473	44.9680
3	21.3495	7.0279	18.3020	34.2074
4	31.9616	7.0698	16.0436	28.9766
5	19.4623	0.8091	10.2601	42.9229
6	12.2804	13.9385	19.0571	36.6827
7	21.0508	11.0130	17.1826	32.7229
8	19.5074	10.9680	16.6510	34.5668
9	26.2808	6.9112	18.8472	28.7695
10	26.1466	9.6041	12.4234	36.6952

Plant	A Nicotine	B Nicotine	C Nicotine	D Nicotine
1	10.0655	8.5977	6.7865	9.9553
2	9.4712	8.1299	10.9249	5.8495
3	9.1246	11.3401	11.3878	10.3005
4	11.3652	9.3470	9.7022	9.7140
5	11.3976	9.3049	8.0371	10.7543
6	11.2936	10.0193	10.7187	8.0262
7	10.6805	9.5843	11.2352	13.1326
8	8.1280	6.4603	7.7079	11.8559
9	10.5066	8.2589	7.5653	11.3345
10	10.6579	5.0106	9.0922	10.4763

   (a) Perform a one-way analysis of variance to test whether there is a significant difference in the average leaf sizes under the four growing conditions. Use $\alpha = 0.05$.

   (b) What conclusions can you reach concerning the effect of growing conditions on the average leaf size?

(c) Perform a one-way analysis of variance to test whether there is a significant difference in the average nicotine contents under the four growing conditions. Use $\alpha = 0.05$.

(d) What conclusions can you reach concerning the effect of growing conditions on the average nicotine content?

(e) Based on the conclusions you reached in parts (b) and (d), do you think the testimony of the tobacco companies' scientists is supported by this experiment? Justify your conclusions.

(f) If you observe significant F-tests in either parts (b) or (d). Carry out the pair-wise tests to determine which mean(s) are different.

(g) Do the nicotine content data suggest violations of the AOV conditions? If you determine that the conditions are not met, perform an alternative analysis, and compare your results to the earlier analysis. Additionally if you observe a significant test, do the pair-wise comparisons using a non-parametric procedure with a Bonferroni correction.

---

2. Game meats, including those from white-tailed deer and eastern gray squirrels, are used as food by families, hunters, and other individuals for health, cultural, or personal reasons. A study by David Holben assessed the selenium content of meat from free-roaming white-tailed deer (venison - VEN) and gray squirrel (squirrel - SQU) obtained from a low selenium region of the United States. These selenium content values were also compared to those of beef produced within (RRB) and outside the same region (NRB). We want to know if the selenium levels are different among the four meat groups.

   The data is found in selenium.csv.

   We want to answer the following questions.

   1. Does meat from larger animals have a higher selenium concentration than small?
   2. Does meat the animals raised in region differ than those raised outside the region?
   3. Does meat from beef differ from non-domesticated animals?

   To answer the above questions, do the following:

   (a) First, conduct the analysis of variance on this data set.

   (b) Specifically identify the null and alternative hypothesis statements for the contrasts that we are testing in the above questions.

   (c) Using Bonferroni's Method, find the contrast estimates, along with their 95% simultaneous confidence intervals and the adjusted p-values for the above research questions. This can be done using a single contrast matrix in R. Display the results, and provide plots of the confidence intervals.

   (d) Don't forget to answer the original research questions using part (c).

   (e) Finally, use Tukey's method to conduct a post-hoc analysis of all of the pairwise contrast estimates, their 95% confidence intervals. Display the results, and provide a plot of the confidence intervals. Comment on any interesting results that you see.

---

3. (You don't have the raw data so you can't use R functions to do this problem automatically for you. Show all of your work.) A researcher collected daffodils from four sides of a building and from an open area nearby. She wondered whether the average stem length of a daffodil depends on the side of the building on which it is growing. Summary statistics are given in the following table. The ANOVA SS(between) is 871.408 and the SS(within) is 3588.54.

	North	East	South	West	Open
Mean	41.4	43.8	46.5	43.2	35.5
SD	9.3	6.1	6.6	10.4	4.7
n	13	13	13	13	13

(a) Create an ANOVA table based upon the information above. Be sure to fully conduct the ANOVA F-test.

(b) Define a contrast to compare the stem length for daffodils from the open area versus the average of the north, south, east and west sides of the building.

(c) Consider applying a (nondirectional) t test to the contrast is part (a). What is the value of t_s and how many degrees of freedom are there?

(d) Determine the P-value. If $\alpha = 0.05$, state your conclusion regardin the contrast in the context of this setting.