Statistics 411/511
Homework 6
Due Tuesday, November 16 by midnight

- **Instructions:** Please see the end of the syllabus for guidelines. Upload your homework to Gradescope via Canvas (access specific homework assignments from the "Assignments" link at the left of the Canvas course page). Your file must be a pdf document. **There will be a one-point deduction if you don't assign pages** (see this Gradescope help video).
- Do the computational part of the homework shortly after completing the week's lab activity.
- The problems are assigned from the **third edition** of the textbook. If you have another edition, consult the copy on one-hour reserve at the library website for the homework problems.
- **Academic Integrity** You are encouraged to *discuss* the homework with other students, but what you turn in must be your own work in your own words. **DO NOT** copy someone else's homework. You may share ideas and R code, but do not share R output or written language. The syllabus contains details and links to OSU's Student Conduct Code and procedure for reporting suspected academic misconduct.

The purpose of this homework assignment is to give you the opportunity to become familiar with the structure of the ANOVA table for a one-way ANOVA.

1. Parts (a)–(e) of this question are to give the numbers that belong in the corresponding boxes of the incomplete ANOVA table below.

Source	df	Sum of Squares	Mean Squares	F-statistic	p-value
Treatment	(a)	(b)	(c)	(d)	0.0001436
Residuals	64	3853.1	(e)		
Total	68	5466.818			

2. For this question, we will work with the data from Griffin, A. M., and W. G. McCarten. "Some Methods for the Quantitative Study of Entozoic Amoebae in Cultures." The Journal of Parasitology 35, no. 2 (1949): 193-98.

   The researchers innoculated each of 50 test tubes with 200,000 individual amoebas. The test tubes contained starch to feed the amoebas. They randomized the 50 test tubes to five treatments:

None	no treatment
Heat10	70 deg heat for 10 minutes
Form10	addition of 10% formalin
FormHeat	heat followed by formalin
HeatForm	formalin followed by heat

   At the end of the 3-day experiment, the yield of amoebas (in units of 1,000 individuals) in each test tube was recorded. The file amoeba.csv is available on Canvas and contains the data.

(Problem 2 continued on next page)

(a) State the name of the response variable.

(b) Use R to produce an ANOVA table for the data. Submit your R code and output.

(c) State the full and reduced models compared by the F-statistic in the ANOVA table in part (b). You can either use the notation introduced on page 17 of Outline 5 or nicknames designated by the textbook.

(d) State the following numbers from the ANOVA table in part (b):

i Extra sum of squares
ii Extra degrees of freedom
iii Residual sum of squares for the full model in part (c).
iv Residual degrees of freedom for the full model in part (c).

(e) Fit the equal means model and obtain an ANOVA table. The code is the same as in item (b), except use formula `Yield~1`. (That's a number "one" on the right-hand side.) Submit your R code and output.

(f) From the ANOVA table in part (e), state the following numbers:

i Residual sum of squares for the reduced model in part (c)
ii Residual degrees of freedom for the reduced model in part (c)

(g) Confirm that the extra sum of squares in part (d) is equal to the difference between the residual sum of squares for the reduced model and the residual sum of squares for the full model. Just use R as a calculator here, and submit your command and result.

(h) State how many mean parameters the full model in part (c) has, and how many mean parameters the reduced model in part (c) has. Confirm that the extra degrees of freedom in part (d) is equal to the difference in number of mean parameters between the full and reduced models. Also confirm that this difference is the same as the difference between the residual degrees of freedom for the reduced model and the residual degrees of freedom for the full model. Again, just use R as a calculator here, and submit your commands and results.