CPaT Stats Lab - Week 5: ANOVA practice

Name: _____
Name: _____
Name: _____

You should be able to finish this lab before noon today, but don't hand it in. At 11:30, I'll post the 'key'…. I want you to have it to use as reference for your Take Home. So, keep it until next Tuesday, and turn in with your takehome midterm.

Due Week 6, hardcopy in Lab next week (Tuesday).

Please delete everything except the questions with your answers inserted in **BOLD** prior to handing in your final assignment as a hardcopy. Bring this completed lab assignment to class next week to hand in. Please be sure your name(s) is (are) on your document.

Upload the associated Excel file to the fileshare:
cpt/Workspace/_StatsLabReports/Week_5/
CushingSkomraWeiss_Wk5statsLab.xlsx (or .xls)
where you type your last name(s) in place of "CushingSkomraWeiss"

Just to be sure that you know how to run an ANOVA in both resampling and JMP, I've prepared a simple example for you. We will do these calculations by hand, then in JMP, and finally using resampling.

Part 1: A "simple as possible" ANOVA – Comparing 3 means – by hand!

The data to use will be on our fileshare in
Handouts\Stats\week 5 - more ANOVA,
in the file wk5-stats-Lab-ANOVA-demo.xlsx
in the Worksheet: 1-way ANOVA-example

Fill in as many of the blanks in the tables below, using your by-hand calculations.

Recall (from Lecture Monday!):

Q1: Fill in the missing cells (???) in the table below
Example: Effects of early snowmelt on alpine plant growth (Delphimiums)

	Unmanipulated	Control	Treatment		
	10	9	12	N groups	???
	12	11	13		
	12	11	15		
	13	12	16		
Ybar-i	???	???	???	Ybar	???
n-i	???	???	???	N samples	???
	???	???	???	**SS amon**	???

				g	

Q2: What kinds of data do you have? _____

What are your null and alternative hypotheses?

H0: _____

Ha: _____

Now do ANOVA calculations.

SS-among: _____

SS-Total: _____

SS-within: _____

F-statistic: _____

Q3: ANOVA Summary of Fit

Rsquare	

Analysis of Variance

Source	**DF**	**Sum of Squares**	**Mean Square**	**F Ratio**	**Prob > F**
Age Class					
Error					
C. Total					

Part 2: A "simple as possible" parametric ANOVA – Comparing 3 means – using JMP!

To see how to run an ANOVA in JMP, consult the JMP documentation, or Judy's JMPCheatSheet in Handouts\Stats.

Use the worksheet 1-way ANOVA-forJMP for this part of the lab.

Q4. Run Shapiro-Wilks tests for each group. Which are likely taken from normal populations and which not? Fill out the table below:

	Group Name	**W value**	**P value*** **	**Normal?**
Group1	1.			
Group2	2.			
Group3	3.			

Q5. Now check if the samples have equal variances.

- What is the Levene's F statistic? _____
- What is the p-value***? _____
- Do the three samples have equal variances? _____

Q6. What are your null and alternative hypotheses?

H0:

Ha:

Now, for comparison with your hand calculations (and to practice using JMP for ANOVA), run a one-way ANOVA to test your null hypothesis.

Q7. Report your ANOVA results by filling in the blanks in the table below:

ANOVA Summary of Fit

Rsquare	

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Age Class					
Error					
C. Total					

Q8. How does your table from Part I compare with the table above?

Q9. Interpret your results.

Q10. Which (if any) groups look like they differ from one another?

in JMP run a Tukey's HSD test to determine which age classes are significantly different from the others.

Q11. Report your F(x,x) = X.XX, p = X.XXXX (an F with x,x degrees of freedom = X.XX, p = X.XXX) and Tukey's lowercase letters on the graph you will make in Q10 below.

Q12. Create a vertical bar graph of to show your results. Insert your graph below.

Q13 Please interpret your results as you would in a scientific paper.

Part 3: Resampling ANOVA.

You will be using the same data for Part 3 as you did for Parts 1,2, but the data need to be reformatted it to make it easier for you to do a resampling ANOVA: see the worksheet 1-way ANOVA-resampling

To analyze these data using Resampling Stats you will first need to calculate a few things. First, recall how many groups we have. # groups is *a* (below).

Then, recall how the sample size is for each group; this is *n*.

In Excel, please calculate the following (these are reported in the table in Q14 below.

a. *Number of groups (a)* _____

b. *sample size for each group/treatment (n)* (report n_i's for i = 1,a if sample sizes are not all the same)

c. *overall mean* (*Ybar)*

d. *average for each group (*$Ybar_i$ *where i=1,a)*

e. *Total Sum of Squares (*SS_{total}*)*

f. *SS Among (*SS_{among}*)*

Q14. Fill in the following table with your calculations:

What is your (actual) SSamong? _____

Number of groups (a)

sample size for each group (n

overall mean (*Ybar)*

average for each group ($Ybar_i$ *where i=1,a)* Group 1:
Group 2:
Group 3:

Total Sum of Squares (SS_{total}*)*

SS Among (SS_{among}*)*

Now, run your resampling ANOVA.

Q15. *What is your Resampling Stats p-value? Is it comparable to your ANOVA p-value from JMP? Speculate as to why or why not.*

Part 4: Optional. If I have time, and if anyone is interested, I have posted another ANOVA dataset for those want more practice: wk5-optional-SpongeData. The dataset is in a plain text file, and your job is to put it into excel in the appropriate format. The four groups you will be comparing are unmanipulated, foam, *Haliclona*, *Tedania*, and the y-variable will be root growth rate.

For this data set, perform the appropriate kind of ANOVA (remember, you can run resampling stats for a data set that meets the parametric ANOVA assumptions, but not vice versa). The data set can be found in the file

Q16: Report the results of this test.

Next week, we might try an n-way ANOVA.

