STATISTICS 402 – Assignment 10
Solution

1. What is a picnic without ants? A student in a statistics class in Australia designed an experiment to see if different types of sandwiches attract different numbers of meat ants (a scavenger ant found throughout Australia). The results were reported by Margaret Mackasack (1994), "Favourite Experiments: An Addendum to What is the Use of Experiments Conducted by Statistics Students?" in the *Journal of Statistics Education*, v2n1. The experiment involved different combinations of bread, filling and butter to make a sandwich. The sandwich was then left next to a meat ant hill. After 5 minutes the experimenter counted the number of ants within a specified radius of the sandwich. The factors of interest were; type of bread (rye, whole meal, multi-grain, or white); the type of filling (Vegemite, Peanut Butter or Ham and Pickle, and whether butter was put on the bread (yes or no). Two sandwiches were made for each combination and the order in which sandwiches were placed near the meat ant hill was completely randomized. Each sandwich was cleaned up and the ants were allowed to return to the hill before placing another sandwich. Below are the data.

		Filling					
		Vegemite		Peanut Butter		Ham and Pickle	
Bread	Butter	Yes	No	Yes	No	Yes	No
Rye		22	18	27	43	68	44
		45	31	50	36	65	54
Whole meal		57	29	42	59	58	34
		42	21	36	47	77	65
Multi-grain		26	42	60	22	63	36
		28	38	47	19	76	59
White		52	42	57	24	66	48
		40	25	51	21	59	53

   a. Use the sample size tables from earlier in the semester to determine the size of the difference in mean quality that can be detected when comparing the:
      - Two levels of butter. Alpha=0.05 and Beta=0.10

        Each level has 24 values. You can detect a difference of between 0.9 and 1.0 standard deviation.

      - Three levels of filling. Alpha=0.05 and Beta=0.10

        Each level has 16 values. You can detect a difference of between 1.2 and 1.4 standard deviations.

      - Four levels of bread. Alpha=0.05 and Beta=0.10

        Each level has 12 values. You can detect a difference of between 1.6 and 1.8 standard deviations.

b. Construct a **complete** ANOVA table (you can, and should, use JMP to do the calculations).

Source	DF	Sum of Squares	Mean Square	F Ratio	Prob > F
Bread	**3**	**196.1667**	**65.3889**	**0.6788**	**0.5736**
Filling	**2**	**4629.8750**	**2314.9375**	**24.0305**	**<.0001**
Bread*Filling	**6**	**273.9583**	**45.6597**	**0.4740**	**0.8208**
Butter	**1**	**1925.3333**	**1925.3333**	**19.9862**	**0.0002**
Bread*Butter	**3**	**195.5000**	**65.1667**	**0.6765**	**0.5749**
Filling*Butter	**2**	**167.0417**	**83.5208**	**0.8670**	**0.4330**
Bread*Filling*Butter	**6**	**2255.1250**	**375.8542**	**3.9016**	**0.0073**
Error	**24**	**2312.0000**	**96.3333**		
C. Total	**47**	**11955.0000**			

c. What factors and/or interactions are statistically significant at the 1% level?

- **Filling; F = 24.03, P-value < 0.0001**
- **Butter: F = 19.99, P-value = 0.0002**
- **Bread*Filling*Butter; F = 3.90, P-value = 0.0073**

d. Compute means for those factors that are statistically significant. Indicate which means are significantly different for the significant factors. Be sure to include the value of the LSD or HSD used.

- **Filling**

HSD = 2.49729(3.47011) = 8.666

Level			Mean
HamPickles	**A**		**57.8125**
Peanut Butter		**B**	**40.0625**
Vegemite		**B**	**34.8750**

Levels not connected by same letter are significantly different.

- **Butter**

LSD = 2.0639(2.8333) = 5.848

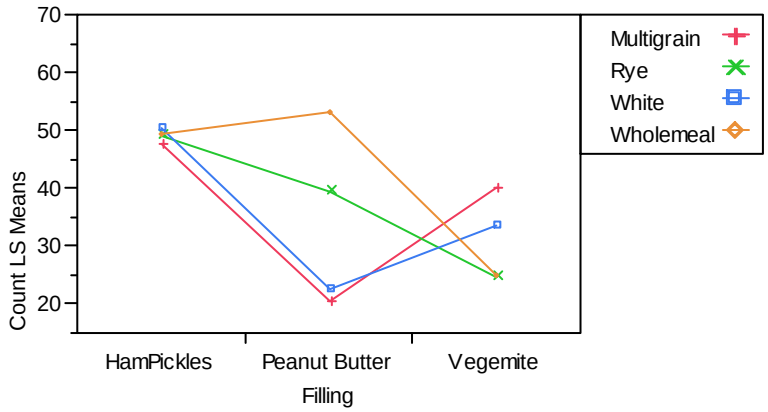
Level			Mean
Yes	**A**		**50.583333**
No		**B**	**37.916667**

Levels not connected by same letter are significantly different.

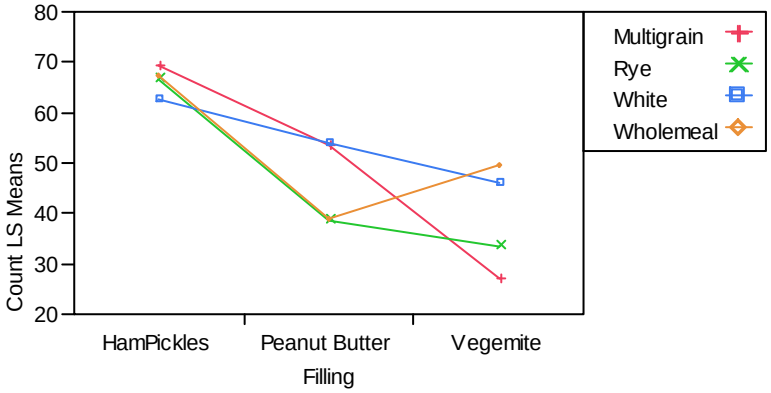
e. Compute means for the significant interactions. Construct and interpret interaction plots for the significant interactions.

		Filling					
		Vegemite		Peanut Butter		Ham and Pickle	
Bread	Butter	Yes	No	Yes	No	Yes	No
Rye		**33.5**	**24.5**	**38.5**	**39.5**	**66.5**	**49.0**
Whole meal		**49.5**	**25.0**	**39.0**	**53.0**	**67.5**	**49.5**
Multi-grain		**27.0**	**40.0**	**53.5**	**20.5**	**69.5**	**47.5**
White		**46.0**	**33.5**	**54.0**	**22.5**	**62.5**	**50.5**

Bread*Filling (No Butter)



Bread*Filling (Yes Butter)



Ham and Pickle tends to have the highest counts of meat ants, higher with butter than without. For multigrain and white bread, peanut butter with butter has fairly high counts but without butter has low counts. For rye and whole-meal, peanut butter with butter has low counts but without butter has high counts. For Vegemite white and rye bread are similar with and without butter. Multigrain with Vegemite and whole-meal with Vegemite act differently with and without butter.

f. Based on your analysis, what sandwich attracts the most meat ants? The fewest meat ants? Support your answer statistically.

Ham and Pickle on any bread with butter tends to attract the most meat ants on average. However, there are other combinations that are not statistically different. Peanut Butter on white or multigrain without butter tends to attract the fewest meat ants on average. However, there are other combinations that are not statistically different. There is no clear winner or loser.

2. An experiment is conducted to explore the relationship between height of step (5.75 in or 11.5 in) and rate of stepping (14 steps/min, 21 steps/min or 28 steps/min) on the heart rate of college students. Six college students were used in the study. There are 6 combinations of step height and stepping rate. Each student experienced each combination. The order was randomized for each student and enough time separated the trials so that students' heart rates could return to a resting rate. The resting heart rate for each student is taken before each trial and the heart rate at the end of the 3 minutes of the stepping regimen is also measured. Below are the data. The values are the change in heart rate from resting to after stepping.

		Stepping Frequency					
		14 steps/min		21 steps/min		28 steps/min	
Student	Step Height	5.75 in	11.5 in	5.75 in	11.5 in	5.75 in	11.5 in
1		16	39	21	52	24	66
2		9	33	21	50	42	60
3		8	18	24	27	27	51
4		9	15	15	28	18	40
5		16	26	28	49	43	59
6		4	15	15	16	20	40

a. Construct a **complete** ANOVA table for the analysis of the change in heart rate. You can use JMP to do the calculations.

Source	**df**	**Sum of Squares**	**Mean Square**	**F**	**Prob>F**
Step Height	**1**	**2916.00**	**2916.00**	**73.84**	**< 0.0001**
Step Freq	**2**	**3314.00**	**1657.00**	**41.96**	**< 0.0001**
Height*Freq	**2**	**152.67**	**76.34**	**1.93**	**0.1658**
Student	**5**	**2114.00**	**422.80**		
Error	**25**	**987.33**	**39.49**		
C. Total	**35**	**9484.00**			

b. What, if any, factors and interactions are significant? Support your answers statistically.

Step Height is statistically significant. With an F = 73.84 and associated P-value of < 0.0001, the low P-value indicates a statistically significant factor.

Step Frequency is statistically significant. With and F = 41.96 and associated P-value of < 0.0001, the low P-value indicates a statistically significant factor.

c. If a factor is significant, what levels are significantly different from other levels? Be sure to include the value of the LSD or HSD used.

Step Height	Mean	Step Frequency	Mean	
11.50	38.00	14 steps/min	17.33	A
5.75	20.00	21 steps/min	28.83	B
		28 steps/min	40.83	C
		LSD = 5.284		

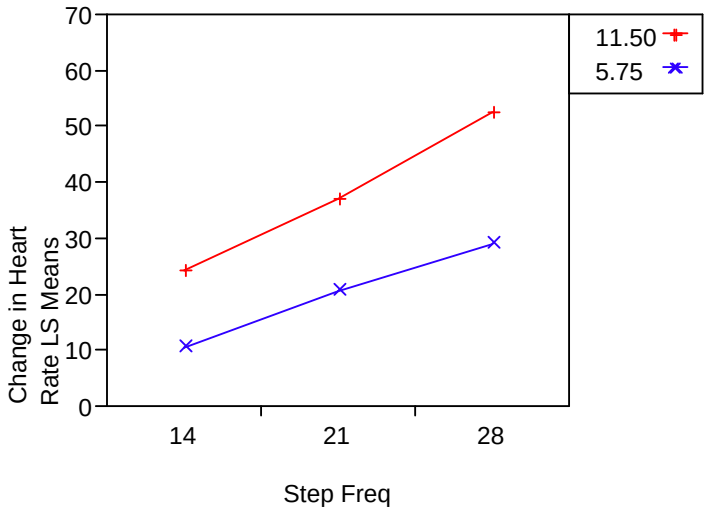
Because there are only two levels of Step Height, the 11.50 inch step height has a statistically higher mean change in heart rate than the 5.75 inch step height.

Stepping Frequency has three levels so we need a value for the LSD. The LSD = 2.05954*2.56558 = 5.284. All three levels of Stepping Frequency are different from each other. 14 steps/min has the lowest mean increase in heart rate, followed by 21 steps/min, and 28 steps/min, which has the highest mean increase in heart rate.

d. Construct and interpret an appropriate interaction plot.

Step Height	14 steps/min	21 steps/min	28 steps/min
5.75 inches	10.33	20.67	29.00
11.5 inches	24.33	37.00	52.67

As Stepping Frequency increases, the change in heart rate increases at about the same rate for each Step Height. There is a slight tendency that as Stepping Frequency increases, the difference between the low and the high Step Height becomes more pronounce. However, the virtually parallel lines indicate that there is very little interaction.



e. Based on your analysis, comment on the effects of step height and rate of stepping on the change in heart rate.

As you increase the Stepping Frequency, heart rate increases by about 12 (on average) for each additional 7 steps/min. The higher Step Height about doubles the change in heart rate, and this is consistent across all Stepping Frequencies.

f. How important was blocking on the student in this experiment? Support your answer with a statistical argument. This may involve additional analysis.

One way to evaluate the effectiveness of blocking is to analyze the data as if it were a completely randomized design and see how the Mean Square Error and conclusions change.

Source	df	Sum of Squares	Mean Square	F	Prob>F
Step Height	**1**	**2916.00**	**2916.00**	**28.21**	**< 0.0001**
Step Freq	**2**	**3314.00**	**1657.00**	**16.03**	**< 0.0001**
Height*Freq	**2**	**152.67**	**76.33**	**0.74**	**0.4864**
Error	**30**	**3101.33**	**103.38**		
C. Total	**35**	**9484.00**			

Notice that the Mean Square Error is now 103.38, over 2½ times bigger than the previous Mean Square Error. This is a big change in the Mean Square Error, however, the analysis and what is statistically significant does not change. Step Height and Stepping Frequency are still statistically significant. The LSD for Stepping Frequency is 2.04227*4.1509 = 8.48, which is substantially larger than before but still results in all levels of Stepping Frequency being statistically different from all other levels.

Therefore, the Mean Square Error is much lower with the block design but the conclusions are the same.

g. If the response had been pulse rate after stepping rather than change in pulse rate, would blocking on student have been more or less important? Explain your reasoning.

It would probably have been more important. By taking change in pulse rate, you are already accounting for differences in the resting pulse rates for the students in the study. If one uses just pulse rate after stepping as the response, these natural differences in resting pulse rates are hidden in the response. By blocking on student the natural differences among students' resting pulse rates are accounted for in the student effect.