STATISTICS 402 – Assignment 8 - Solution

1. For each of the following problem statements, identify the response, conditions and experimental material. Indicate the design (completely randomized, randomized complete block, Latin square) that is used to collect the data and indicate if factorial crossing is used. Construct a partial ANOVA table by identifying all sources of variability and associated degrees of freedom.

    a. Industrial psychologists wish to investigate the effect of music in the factory on the productivity of workers. Four distinct music programs (Country, Soft Rock, Hard Rock, and Classical) and no music make up the five treatments. In order to account for variability from day to day and week to week, a different program is used on each day (M, T, W, R, F). The programs are rotated from week to week, so that over 5 randomly chosen weeks each program appears once on any specific working day of the week.

    Response: Productivity
    Conditions: Music program, Day, and Week
    Experimental Material: Workers

    Latin Square Design with music program as the factor of interest. Day and week are nuisance factors.

Source	df
Music Program	4
Day	4
Week	4
Error	12
C. Total	24

    b. An experiment is conducted to compare the energy requirements of three physical activities: running, walking, and bicycle riding. Eight subjects are asked to run, walk and bicycle a measured distance, and the number of kilocalories expended per kilometer is measured for each subject during each activity. The order of the activities is randomized for each individual with time for recovery between activities.

    Response: Energy (kilocalories/kilometer)
    Conditions: Physical activity (running, walking, and bicycle riding)
    Experimental Material: Subjects

    Randomized Complete Block Design with physical activity as the factor of interest and subjects as the blocks.

Source	df
Physical Activity	**2**
Subject	**7**
Error	**14**
C. Total	**23**

c. An experiment is performed to see whether different operators obtain different results in the routine analysis to determine the amount of nitrogen in soil. 50 soil samples are chosen at random and divided at random into 5 groups of 10. Each operator is assigned a group of 10 soil samples at random and asked to determine the amount of nitrogen in each sample.

Response: Amount of nitrogen
Conditions: Operators
Experimental Material: Soil samples

Completely Randomized Design with Operator as the factor of interest.

Source	df
Operator	**4**
Error	**45**
C. Total	**49**

d. A sports physiologist is conducting an experiment on eye focus time. She is interested in the distance of the object from the eye on the focus time. Four different distances are of interest. She chooses 10 players at random from a college softball team. Each player will have her focus time measured at each of the four distances. The order of the distances will be randomized for each player.

Response: Eye focus time
Conditions: Distance
Experimental Material: Soft ball players

Randomized Complete Block Design with distance as the factor of interest and players as blocks.

Source	df
Distance	**3**
Player	**9**
Error	**27**
C. Total	**39**

e. A mechanical engineer is studying the thrust force developed by a drill press. She is interested in the effects if drill speed and drill size. She has five drill speeds and two different drill sizes. A single operator drills 4 holes using each of the speed and size combinations in pieces of aluminum. The order in which the 40 pieces of aluminum are chosen and holes are drilled is randomized.

Response: Thrust force
Conditions: Combination of drill speed and drill size
Experimental Material: Pieces of aluminum

Completely randomized design with factorial crossing of drill speed and drill size.

Source	**df**
Drill speed	**4**
Drill size	**1**
Speed*Size	**4**
Error	**30**
C. Total	**39**

2. The following is adapted from a problem that appears on page 173 of *Design and Analysis of Experiments*, 3rd edition by Douglas Montgomery.

The effect of five different ingredients (A, B, C, D, and E) on the reaction time of a chemical process is being studied. Each batch of new material is only large enough to permit five runs to be made. Furthermore, each run requires approximately 1½ hours, so only five runs can be made in one day. The experimenter decides to run the experiment as a Latin square so that day and batch may be accounted for in the analysis. Be sure to turn in appropriate JMP output with your assignment.

	Day				
Batch	1	2	3	4	5
1	A=9	B=7	D=2	C=8	E=2
2	C=12	E=2	A=8	D=3	B=8
3	B=5	A=10	C=11	E=1	D=4
4	D=3	C=9	E=3	B=6	A=11
5	E=4	D=2	B=6	A=9	C=8

a. Compute an overall grand sample mean, as well as ingredient, batch and day sample means.

Grand Sample Mean = 6.12

Ingredient	Mean	Batch	Mean	Day	Mean
A	**9.4**	**1**	**5.6**	**1**	**6.6**
B	**6.4**	**2**	**6.6**	**2**	**6.0**
C	**9.6**	**3**	**6.2**	**3**	**6.0**
D	**2.8**	**4**	**6.4**	**4**	**5.4**
E	**2.4**	**5**	**5.8**	**5**	**6.6**

b. Compute the estimated effect for each ingredient. Which ingredients appear to shorten the reaction time?

Ingredient	Mean	Estimated Effect
A	**9.4**	**3.28**
B	**6.4**	**0.28**
C	**9.6**	**3.48**
D	**2.8**	**–3.32**
E	**2.4**	**–3.72**

Ingredients D and E appear to shorten reaction time the most.

c. Construct an ANOVA table including all appropriate sources of variation.

Source	df	SS	MS	F	Prob > F
Ingredient	**4**	**239.04**	**59.76**	**31.017**	**< 0.0001**
Batch	**4**	**3.44**	**0.86**		
Day	**4**	**5.04**	**1.26**		
Error	**12**	**23.12**	**1.927**		
C. Total	**24**	**270.64**			

d. Are there statistically significant differences among the five ingredients in terms of average reaction time? Report the appropriate F- and P-values and explain why these support your answer.

F = 31.017, P-value < 0.0001. The large F value and small P-value indicated that there are statistically significant differences among the sample means for some of the ingredients.

e. If there are statistically significant differences among the ingredients, indicate where those differences lie. Give the value of HSD. Clearly state which ingredients have statistically significant differences in sample means and which do not.

The multiplier for Tukey's HSD is q*=3.18738.

$$\text{HSD} = q^* \sqrt{MSError} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = 3.18738\sqrt{1.9267}\sqrt{\frac{1}{5} + \frac{1}{5}} = 3.18738(0.87788)$$

HSD = 2.798

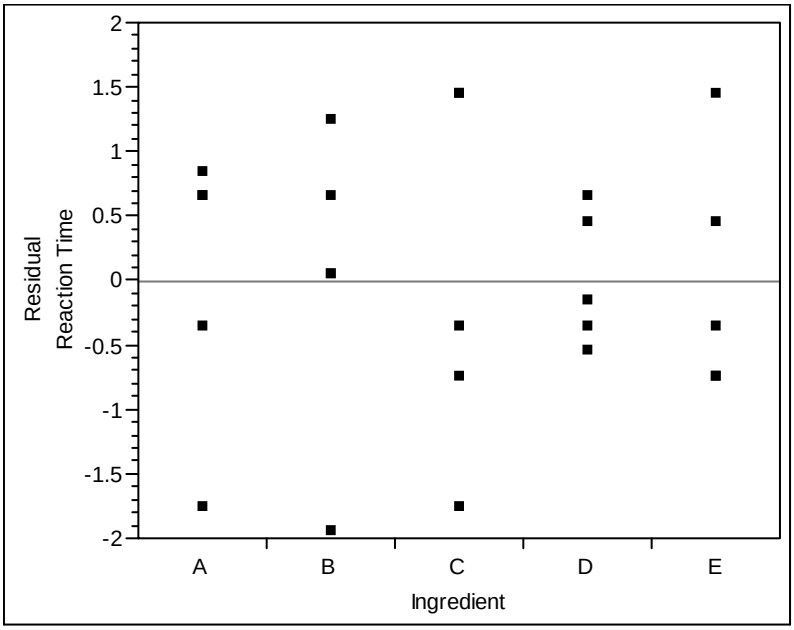
Ingredient	Mean	
C	9.6	A
A	9.4	A
B	6.4	B
D	2.8	C
E	2.4	C

Ingredients with different letters are statistically different.

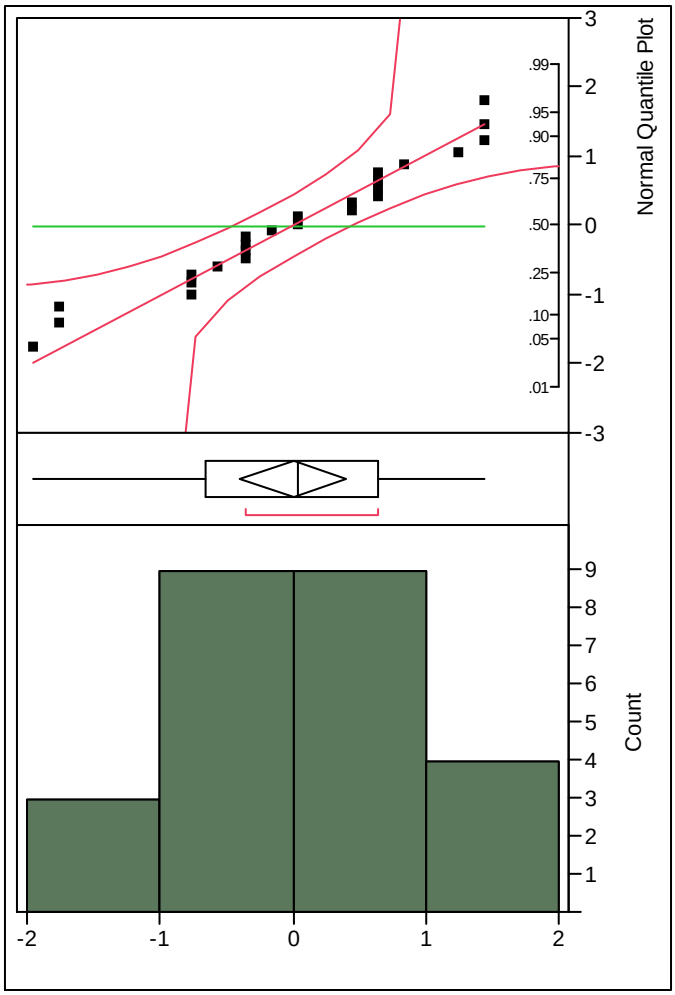
f. If one wishes to reduce reaction time, what statistically valid recommendations can you make for choice of ingredient? Explain your reasoning.

To reduce reaction time, use either ingredient D or E. These two ingredients have the lowest mean reaction times and those times are statistically different from the means of the other ingredients. There is no statistically significant difference in reaction time between the two ingredients, D and E.

g. Look at the residuals and comment on whether the equal standard deviation condition and the normally distributed errors condition are satisfied.



Ingredients B and C have slightly more variation than the other ingredients. However, overall the variation within each ingredient is approximately the same. The equal standard deviation condition is most likely met.



The distribution of residuals is symmetric and mounded in the middle (histogram). The sample mean and sample median are about the same (box plot) also indicating a symmetric distribution. The points on the normal quantile plot follow the normal model line fairly well. The normal distribution condition is probably met.

h. If these data were analyzed as if they came from a completely randomized design, give the ANOVA table with sources of variation, degrees of freedom, sums of squares and mean squares. **Do not use JMP or another computer program to do this part, you should be able to do this from the information in a.**

Source	**df**	**SS**	**MS**	F	**Prob > F**
Ingredient	**4**	**239.04**	**59.76**	**37.82**	**< 0.0001**
Error	**20**	**31.60**	**1.58**		
C. Total	**24**	**270.64**			