Chapter 6 Solutions

6.1 a. The researchers measured the calcium concentration in the plasma of each bird after the birds were treated, so this is the response variable.

b. The two factors were sex and additive given to the birds.

c. Sex was observational (we can't assign birds to a given sex) and has two levels: male and female. Additive was experimental—some birds were randomly assigned to receive the hormone in their diet and the others were given a control diet. Additive had two levels—either the birds got the hormone or they did not.

d. No, this was not a complete block design. In a complete block design, each combination of levels of factors occurs exactly once. In this design, there were five birds of each sex that got the same hormone level.

6.2 a. The researchers dried the plants at the end of the study and measured the weight of what was left. This weight is the response variable.

b. The two factors were the configuration of the plants in the cups (how many of each species was in a cup) and the amount of nutrients that the cups received.

c. Both factors were experimental because both were assigned to the cups. The factor for configuration of the plants had five levels (20 D.i., 15 D.i. and 5 D.s., 10 of each, 5 D.s. and 15 D.i., and 20 D.s.) and the nutrient amount had two levels (normal amount and low amount).

d. No, this was not a complete block design. In a complete block design, each combination of levels of factors occurs exactly once. In this design there were 2 cups for each combination of plant configuration and nutrient amount.

6.3 a. The researchers used a score to measure each subject's adaptation. This score is the response variable.

b. The two factors are the subjects themselves and the shock treatment applied to them.

c. The subject factor was observational (there was no assignment here) and there were 18 levels because there were 18 subjects. The shock treatment applied was experimental (the types of shock treatment were assigned in random order to each subject), and there were three levels of this factor: shock during stuttering, shock after stuttering, no shock.

d. Yes, this is a complete block design. Each subject is a block with each treatment being assigned in a random order to each subject.

6.4 a. The researchers were interested in the math score of the students under different conditions. This is the response variable.

b. The two factors are the hyperactive diagnosis and the amount of noise under which the students attempted the math problems.

c. Hyperactive diagnosis was observational (the researchers could not assign particular children to be hyperactive) and had two levels: hyperactive or not. The amount of noise was experimental since children were assigned to complete the math problems in environments with different amounts of noise. There were two levels to this factor (high noise and low noise).

d. No, this is not a complete block design. In a complete block design, each combination of levels of factors occurs exactly once. While we do not know exactly how many children participated in this experiment, it seems clear that there was more than one child for each combination of hyperactive diagnosis and amount of noise.

6.5 a. The measurement that researchers were interested in analyzing was the running speed of the dogs, so that is the response variable.

b. The two factors are the dogs and the diets.

c. The dogs factor is observational since there is no assignment involved and there are five levels (the number of dogs in the study). The diet factor is experimental since its levels are assigned (in a random order) to each dog. There are three diets, so this factor has three levels.

d. Yes, this is a complete block design. Each dog is a block with each diet being assigned in a random order to each dog.

6.6 a. The researchers were ultimately interested in how much the rats ate, so amount of food eaten is the response variable.

b. The two factors were sex and hormone type.

c. Sex is an observational factor (we can't assign sex to rats) and has two levels. Hormone is an experimental factor (the researchers assigned specific rats to get the different hormones) and has two levels.

d. No, this is not a complete block design. In a complete block design, each combination of levels of factors occurs exactly once. While we do not know exactly how many rats participated in this experiment, it seems clear that there was more than one rat for each combination of sex hormone type.

6.7 a. If the F-ratio is near 1, this is not an unusual occurrence. This suggests that we cannot rule out random variation as being the reason for the difference that we have observed in the copper levels of the various rivers.

b. If the F-ratio is near 10, this is an unusual occurrence. This suggests that it is unlikely that random variation is the reason for the difference that we have observed in the copper levels of the rivers. We would conclude that the average amount of copper in at least one river is different from the average amount of copper in the other rivers.

6.8 a. If the F-ratio is near 0.5, there is virtually no support for the idea that the variance between the groups is larger than the variance within the groups. With an F-ratio near 0.5, the amount of variability within the sample groups is almost double the amount of variability between the groups, and we have no reason to suspect that there are differences between the sleep phases.

b. If the F-ratio is near 1, this is not an unusual occurrence. This suggests that we cannot rule out random variation as being the reason for the difference that we have observed in the heart rates for the different phases of sleep.

6.9 a. The degrees of freedom for each factor is the number of levels for that factor minus 1. In this case, that is $4 - 1 = 3$.

b. The degrees of freedom for each factor is the number of levels for that factor minus 1. In this case, that is $5 - 1 = 4$.

c. The residual degrees of freedom for this model is $(K-1)(J-1)$, which in this case is $(4-1)(5-1) = 12$.

6.10 a. The degrees of freedom for each factor is the number of levels for that factor minus 1. In this case, that is $3 - 1 = 2$.

b. The degrees of freedom for each factor is the number of levels for that factor minus 1. In this case, that is $3 - 1 = 2$.

c. The residual degrees of freedom for this model is $(K-1)(J-1)$, which in this case is $(3-1)(3-1) = 4$.

6.11 a. Large

b. Less

6.12 a. Smaller

b. More

6.13 True. Each block must have all levels of one factor randomly assigned to an experimental unit within the block. If a factor is observational, it cannot be assigned to experimental units.

6.14 True. The errors must still be independent, normally distributed, have mean 0, and have the same variance.

6.15 True. In a complete block design, each block provides one unit for each level of the treatment factor, so all possible combinations of block and treatment are present.

6.16 False. There are three different ways that block designs can occur. Reusing subjects is one of them, but block designs also can be created by subdividing a large plot of land into equal-sized smaller plots with the number of plots being equal to the number of treatments; and by sorting units into equal sized groups of similar individuals with the size of the group equal to the number of treatments.

6.17 a. Version B is a block design. Version A is a completely randomized design, with all 18 units (samples of oil) interchangeable. In Version B, each larger sample is similar to a block: It provides a set of smaller samples (units).

b. The advantage of using blocks comes from having groups of similar units. If there is variation from one container to the next, using only six larger containers instead of 18 smaller ones may reduce variability and give a more powerful test of the effect of aging.

6.18 Blocks by subdividing: RiverIron. In this case, the blocks are the stream locations and the water samples at each site are the units. Blocks by grouping: RadioactiveTwins. In this case, the blocks are the pairs of twins and the subjects are the units. Blocks by reusing: FranticFingers, where the blocks are the subjects and the units are the time periods.

6.19 First reason: Reusing subjects may not work, because the treatment or measurement changes the subject. Once subjects have read one set of results (e.g., "most complied"), you can't present them with a different set of results (e.g., "most refused") and expect them to take you seriously. Once half the leafhoppers in a dish have died, you can't bring them back to life in order to reuse them. Second reason: Reusing subjects may not work, because the treatment or measurement takes too long. A potential example could come from agricultural experiments where it takes an entire growing season to measure the effect of a treatment. In fact, for some forestry experiments, it may take many years to measure the effect of a treatment.

6.20 a. One: In this example, differences between subjects are large. (Subject I and IV were slow; Subject II was fast.) Reusing subjects allows the effect of all three drugs to be measured on the same subject, which makes residual variation much smaller than in the completely randomized design. Two: The block design gets multiple measurements from the same subject. This design gives you 12 response values using only 4 subjects because each subject provides three units.

b. Answers may vary. Three possible reasons are given here. One: The completely randomized design avoids possible carryover effects of the drugs. If the effect of a drug lasts several days, for example, it would not be a good plan to use the same subject on three successive days. Two: The completely randomized design can be carried out all at once in a single day. Reusing subjects, as in the block design, takes several days. Three: The completely randomized design with 12 subjects gives 9 df for error; the block design gives only 6 df for error.

6.21 Answers will vary. One possible answer is given here. One: frantic fingers. A one-way design might look only at the effect of the drug. As illustrated in the text examples, you get a much more informative experiment with an additional factor, subjects crossed with drug. Two: river iron. It would be possible to compare just iron level on the various rivers. By crossing river with a second factor, location on the river, researchers got a much more informative experiment for very little additional cost or effort. Additional possibilities include many of the examples and exercises from Chapter 5. These can be extended by crossing the existing factor with some other factor.

6.22 If your factor of interest is observational, it is not possible to randomly assign levels of the factor to units, and therefore, usually, not possible to create blocks. Note, however, that block designs that are *not* randomized can occur with an observational factor of interest as in certain longitudinal studies. For example, a study follows ten overweight men (subjects) through 12 weeks of a weight-loss program, measuring blood pressure (response) at weeks 0, 3, 6, 9, and 12 (observational factor of interest with five levels). In this case, each subject is a block.

6.23 a. The order of the treatments was randomized for the usual three reasons: Randomizing protects against bias, permits conclusions about cause, and justifies using a probability model. To be more specific and concrete, it is quite possible that there would be a carryover effect from a treatment, which would mean that the order of the treatments has an effect on the response values. Randomizing the order protects against bias from this carryover effect.

b. The main reason for using a block design is to reduce residual variation. Differences from one subject to the next were expected to be large as in the finger tapping study. The block design makes it possible to measure the effects of both treatments and control on the same subject. Put differently, in the block design, between-subject variation is not included in the mean square for error; in the completely randomized design, it is.

6.24 a. If the control attempt at the maze always came before the treatment attempt, it would be impossible to know whether the improved times were due to the treatment or to the effect of practice. Without randomization, inference without cause cannot be justified here.

b. There is more than one reasonable design, but any reasonable design should deal with three different nuisance effects: differences between subjects, differences between mazes, and the effect of order. There is one factor of interest, with two levels: fragrance and control. Because subject differences are expected to be large, subjects should serve as blocks. Because there are only two mazes, each subject can provide only two units, one for each maze. Some subjects will try maze A first, then maze B; the rest will do B first and then A. The assignment of subjects to AB or BA should be randomized. The treatment order should also be randomly assigned, either fragrance first, then control, or control first, then fragrance. In all, then, there are four combinations: A:Control first then B:Fragrance, A:Fragrance first then B:Control, B:Control first then A:Fragrance, and B:Fragrance first then A:Control. One possible design would randomly assign combinations to the 20 subjects without any restrictions on how often each combination occurred. An alternative would still use random assignment but would ensure that all four combinations occurred equally often. To accomplish this, randomly divide the 20 subjects into four groups of 5. Subjects in the first group get combination 1, those in the second group get combination 2, etc.

6.25 Blocks are typically created in one of three ways: by reusing, by subdividing, or by grouping. **Reusing:** A cloud cannot be reused, because it is changed by the treatment. After the rainfall has been measured, the cloud is no longer the same. **Subdividing:** To create blocks by subdividing, you would have to be able to seed half of each cloud, and you would also have to be able to measure rainfall from each half of the cloud. Neither is practical. **Grouping:** To create blocks by grouping you would have to have a good way to classify rain clouds into types. In principle, this might appear to be possible—after all, there is a standard way to do this—but if clouds are grouped, there is no practical way to seed only some of the clouds within a group, so in practice, grouping wouldn't work either. In fact, it turns out that there is only one kind of cloud worth seeding anyway.

6.26 a. The study is experimental because each subject listened to all four lists. For it to be a good experiment, we would expect that randomization was used in assigning the order that the lists were presented to the subjects.

b. There is one factor of interest (list) and one nuisance factor (subject).

c. The experimental units are time slots, four from each subject.

d. Each subject is a block of four units.

6.27 a.

	Complete block	Completely randomized
Diets	2	2
Blocks	4	0
Residual	8	12
Total	14	14

b. The block design uses fewer dogs (5 instead of 15). The block design takes longer because each dog provides three time slots instead of just one. The block design has fewer df for residual (8 instead of 12). The block design splits off between-dog differences from residual error. When these differences are large, the block design gives a more powerful test for the effect of the treatment.

6.28 a. Let each subject provide a block of four time slots. For example, each time slot might be a stretch of six weeks. For each subject, randomly assign one treatment to each time slot, with each subject getting all four treatments, one at a time in a random order, with a separate random order for each subject. An equivalent way to get the same design: There are $4! = 24$ possible orders. List and number them: (1) IFAP, (2) IFPA, ... , (24) PAFI. Randomly assign a number from 1 to 24 to each subject to determine the order for that subject.

b. There are several possible correct answers. Here is one:

	Time			
Subject	1	2	3	4
I	I	F	A	P
II	F	A	P	I
III	A	P	I	F
IV	P	I	F	A

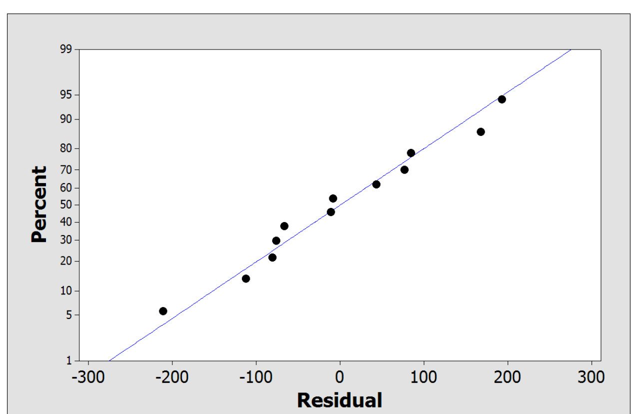
6.29 a. Use a randomized complete block design: For each rat, randomly assign one treatment (leptin or insulin) to the first time slot, with the other hormone assigned to the second time slot. Use a separate randomization for each rat.

b. This is a situation where it is reasonable to expect the treatment to change the subject (rat in this case), and so reusing subjects is not a good idea. First, being injected (with anything) might cause a reaction by the rat, but the second injection might produce a different reaction, or perhaps none at all, as the rat would have some experience with being injected. Also, the first and second time slots are not the same, because the rat is younger and lighter for the first, older and heavier for the second.

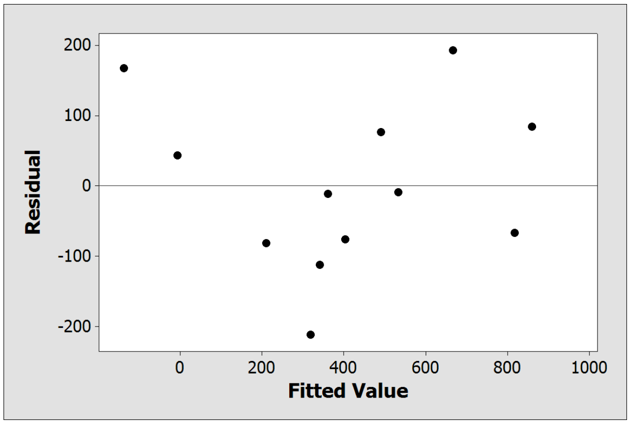
6.30 a.

Source	DF	SS	MS	F	P
River	3	542458	180819	7.04	0.022
Site	2	442395	221198	8.61	0.017
Error	6	154083	25680		
Total	11	1138936			

b. No. The normal probability plot appears to be consistent with the idea that the residuals are normally distributed. There does not appear to be any pronounced departure from linearity in this plot.



c. Yes, there is an indication of a problem with the residuals versus fitted values plot. There is pronounced curvature.

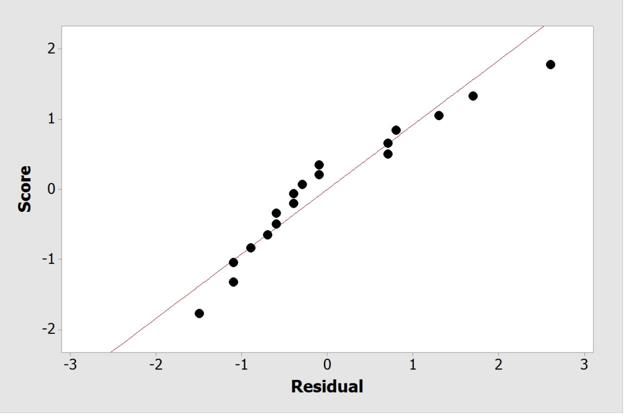


d. The leftmost fitted value appears to be about −200. But the value for the amount of iron cannot be negative. This indicates a major problem with the model.

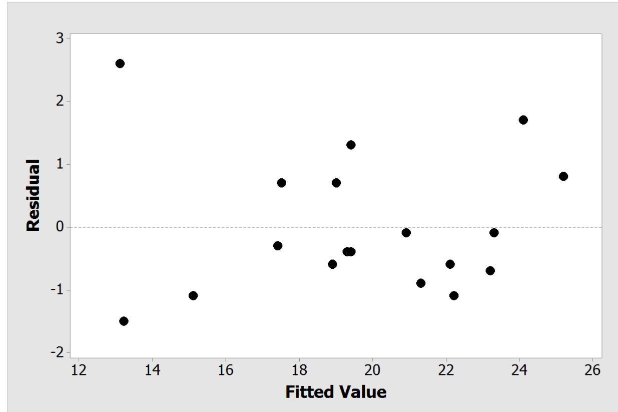
6.31 a.

Source	DF	SS	MS	F	P
Shrew	5	194.160	38.832	19.34	0.000
Phase	2	15.240	7.620	3.79	0.059
Error	10	20.080	2.008		
Total	17	229.480			

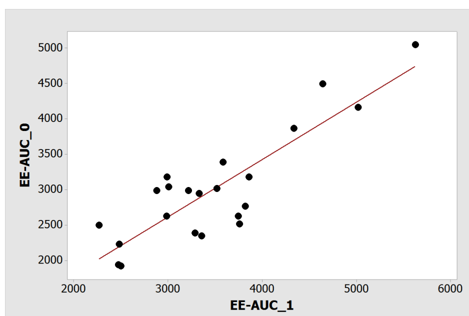
b. No. The normal probability plot appears to be consistent with the idea that the residuals are normally distributed. There does not appear to be any pronounced departure from linearity in this plot.

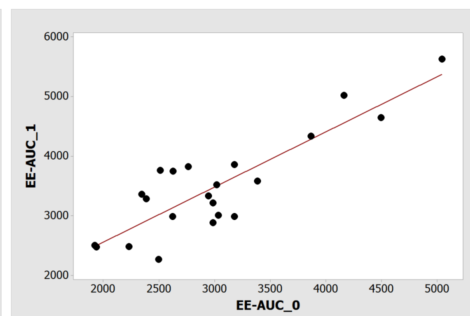


c. Yes, there is an indication of a problem with the residuals versus fitted values plot. There is an outlier in the upper left corner and possible linearity among the rest of the points as a consequence of the outlier.



6.32 a. There are two ways to create this plot—depending on which variable you place on the x-axis. We give both plots below.

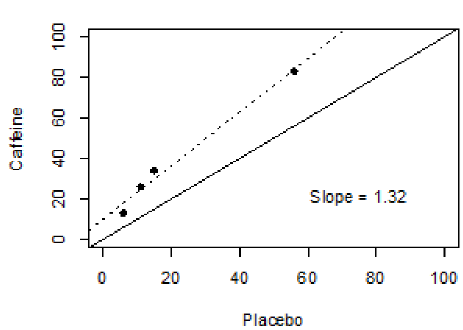


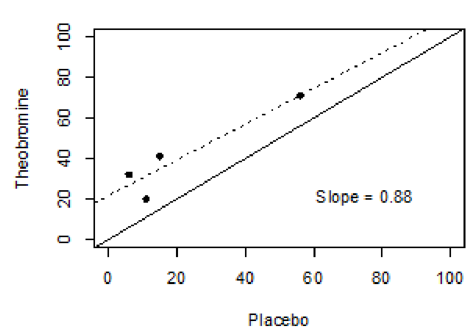


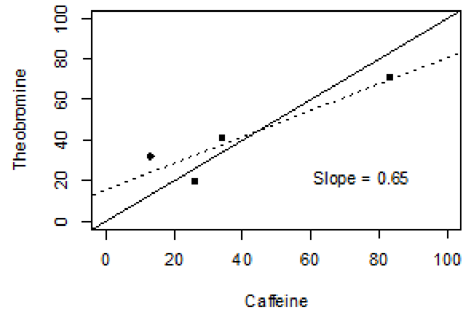
The two slopes given by these graphs are 0.8 and 0.92, both close to 1.

b. Since the slopes in the Anscombe block plot is close to 1, no transformation is necessary.

6.33 a. The figure follows.

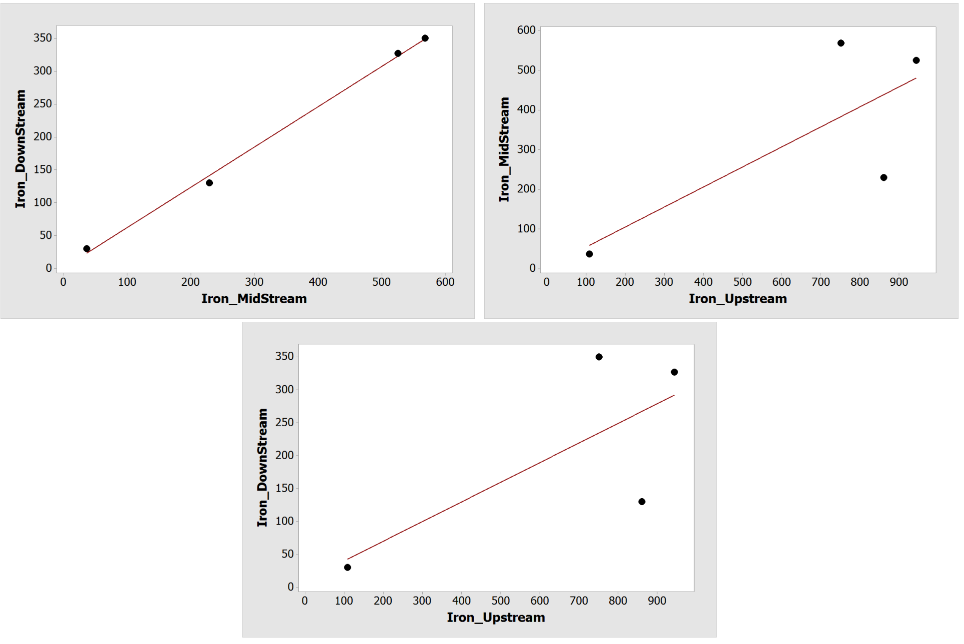






b. For all three plots, the fitted lines have slopes roughly equal to 1, especially if you take into account how few subjects there are. If you allow for unit-to-unit variation, there is no strong evidence that a model of additive main effects should be discarded.

6.34 a. The figure follows.



b. The three slopes are 0.3, 0.5, and 0.6. These are all approximately 0.5. Since they are all different from 1, it suggests that there might be a transformation that would help.

6.35 As the researchers suspected, the food cooked in the iron pots has the most iron in it. Food cooked in clay pots has much less iron than the food cooked in iron pots, but a slightly higher amount of iron than food cooked in aluminum pots. There are, however, two low outliers for the meat and vegetable dishes cooked in the clay pots, which might affect the significance of a difference between clay and aluminum pots. There are also concerns about the spread of the observations. There is much smaller spread for the values of iron in the poultry dish. There is somewhat more spread in the amount of iron in the vegetable dish and quite a bit more spread in the amount of iron in the meat dish. Since the smallest amount of variability is in the observations in the middle of the scale, a change of scale is not likely to help.

6.36 a.

Variable	Row	Mean
Ht4	a	1.160
	b	1.570
	c	1.250
	d	2.260
	e	2.460

Variable	Acid	Mean
Ht4	1.5HCl	1.466
	3.0HCl	1.084
	water	2.670

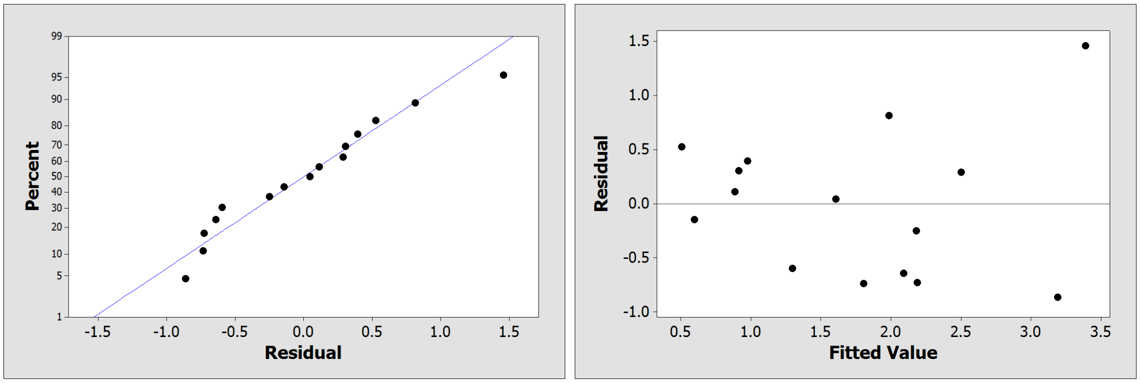
Variable	Mean	StDev
Ht4	1.740	1.105

b. `Two-way ANOVA: Ht4 versus Acid, Row`

Source	DF	SS	MS	F	P
Acid	2	6.8516	3.42578	4.51	0.049
Row	4	4.1826	1.04565	1.38	0.324
Error	8	6.0724	0.75905		
Total	14	17.1066			

`S = 0.8712   R-Sq = 64.50%   R-Sq(adj) = 37.88%`

c. First, we check conditions. The normal probability plot suggests that normality is reasonable for the residuals since it displays a roughly linear pattern, and the residuals versus fits plot shows a relatively even band of points, indicating that the equal variances condition is met.



Source	DF	SS	MS	F	P
Acid	2	6.8516	3.42578	4.51	0.049
Row	4	4.1826	1.04565	1.38	0.324
Error	8	6.0724	0.75905		
Total	14	17.1066			

d. Yes. The P-value for the *Acid* factor is significant, which suggests that the variability observed is not likely to be due to randomness but rather the treatment instead.

e. No. The P-value for the *Row* factor is not significant. It is perfectly reasonable to think that the variability seen among the plants in different rows is due to randomness.

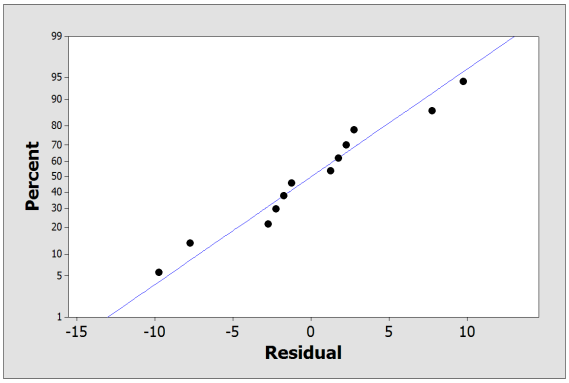
6.37 The $MSE = 0.75905$ and the t^* has 8 df, which means $t^* = 2.306$, so the margin of error for all intervals is $2.306\sqrt{0.75905\left(\frac{1}{5}+\frac{1}{5}\right)} = 1.271$. The intervals then are: for $1.5HCL$–$3.0HCL$: $(1.466-1.084) \pm 1.271 = (-0.889, 1.653)$; for $Water - 1.5HCL$: $2.670 - 1.466 \pm 1.271 = (-0.067, 2.475)$; and for $Water - 3.0HCL$: $2.670 - 1.084 \pm 1.271 = (0.315, 2.857)$. The only interval that does not contain 0 is the one for $Water - 3.0HCL$. Therefore, there is a significant difference in the growth between those plants given water and those given 3.0 HCL.

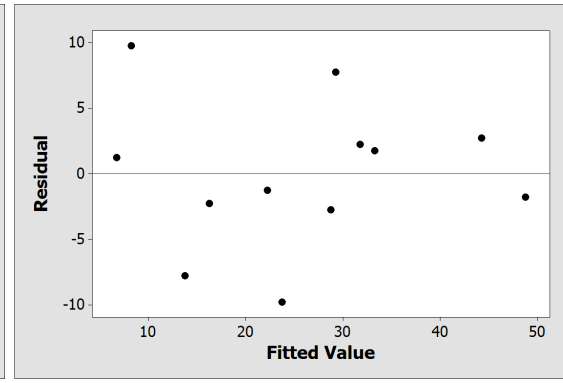
6.38 a. The overall mean number of unpopped kernels is 25.58. The mean number of unpopped kernels for Orville is 17.83 and for Seaway is 33.33. This means that our estimate for α_1 (for Orville) is $17.83 - 25.58 = -7.75$ and our estimate for α_2 is $33.33 - 25.58 = 7.75$.

b.

Source	DF	SS	MS	F	P
Brand	1	720.75	720.750	10.45	0.023
Trial	5	1181.42	236.283	3.43	0.101
Error	5	344.75	68.950		
Total	11	2246.92			

Checking conditions, we look at the normal plot of the residuals and the residuals versus fits plot. Neither plot shows any indication that the model is inappropriate. Note that, while the normal plot has two distinct lines between -5 and 0 and between 0 and 5, the overall pattern to the entire plot is roughly linear.



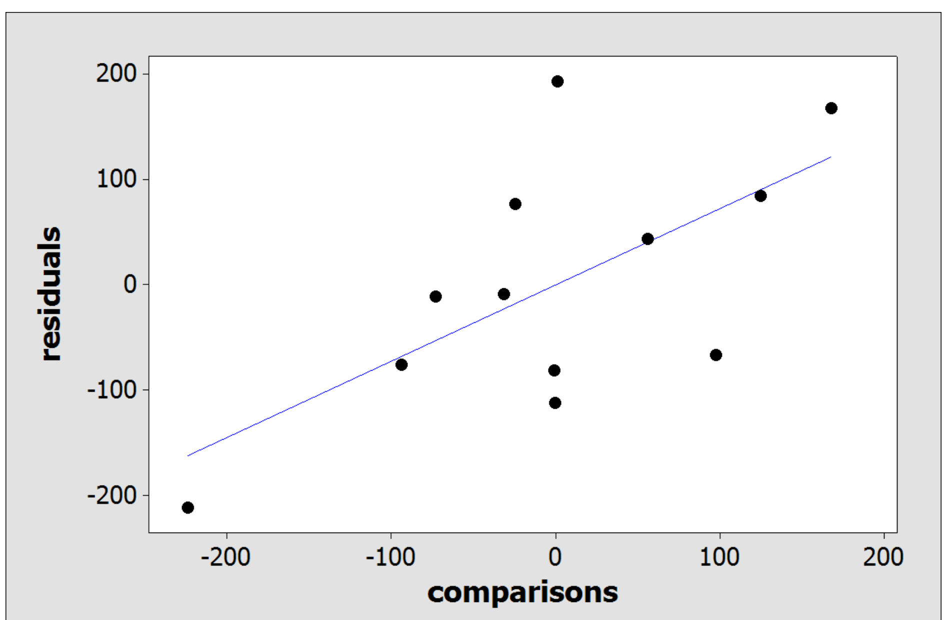


c. The brand does appear to make a difference in the number of unpopped kernels with a P-value of 0.023. It appears that Seaway has more unpopped kernels than Orville. The trial does not appear to make a difference with a P-value of 0.101.

6.39 a. Here is a table of the comparison values.

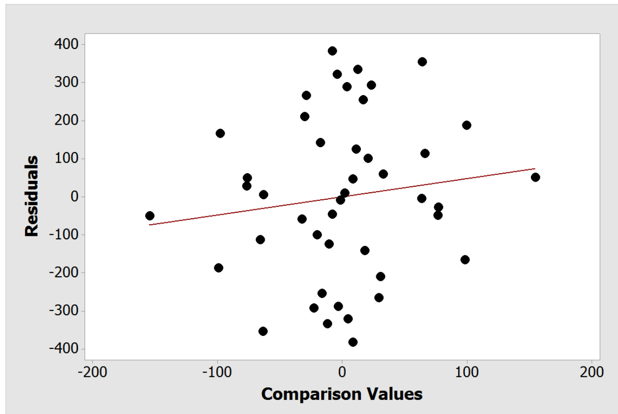
Site	Grasse	Oswegatchie	Raquette	St. Regis
Up	124.93	0.97	−223.54	97.64
Mid	−31.28	−0.24	55.97	−24.45
Down	−93.64	−0.72	167.56	−73.19

b.



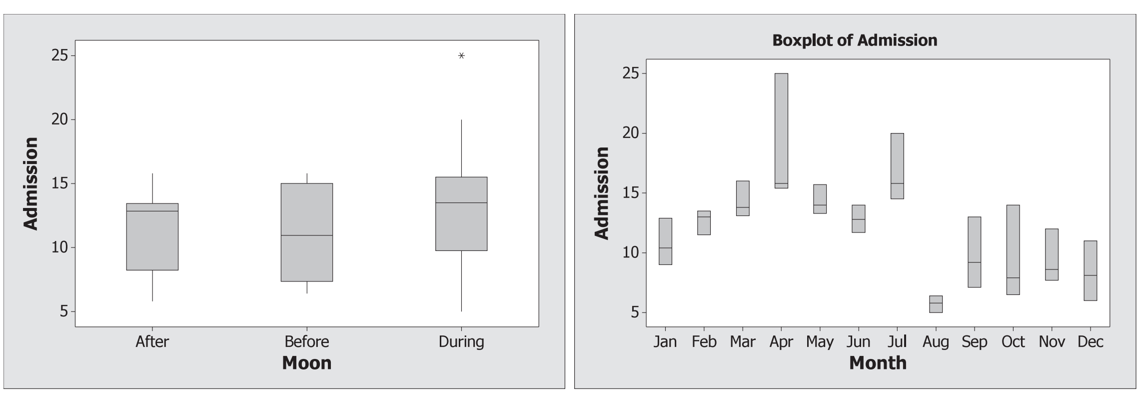
c. The slope appears to be approximately 0.75. This means that $p = 1 - 0.75 = 0.25$ and the transformation would be the fourth root.

6.40 The graph follows.



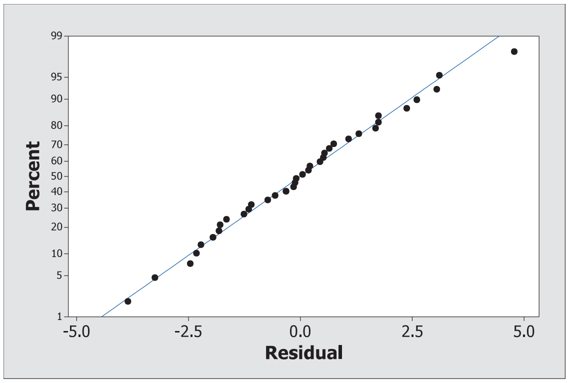
While there is a line indicated (the slope is about 0.5), there is so much variability around the line that there is not a strong suggestion for transformation. Leave the data as it is.

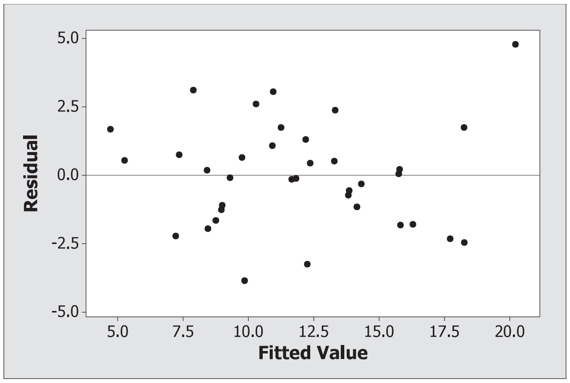
6.41 First, we explore the data to see if an ANOVA model might be useful. That is, do we think that either *Moon*, *Month*, or both, might help explain the number of admissions. We give two boxplots below. In the first, we see that there might be a full moon affect since Q1, the median, and Q3 are all higher for the "during" than the other two. The second boxplot shows that *Month* almost certainly plays a role, with more admissions in the spring and early summer.



Before computing an ANOVA model, we consider the conditions. The data are random and independent to the extent in any given year, once we know the month and the moon, the actual admissions will vary and there is no reason to think that knowing the value of one observation will tell us more about any other observation. We feel more comfortable with this idea since there is a gap of 3 days between the "before full moon," the full moon, and the "after full moon" observations.

The normal plot of the residuals and the residuals versus fits plot also confirm that conditions are met.





Now we can look at the output from the ANOVA procedure itself. It is given below:

Source	DF	SS	MS	F	P
Month	11	455.583	41.4166	7.13	0.000
Moon	2	41.514	20.7569	3.57	0.045
Error	22	127.819	5.8100		
Total	35	624.916			

```
S = 2.410   R-Sq = 79.55%   R-Sq(adj) = 67.46%
```

From this, we see that both $Month$ and $Moon$ are significant, though $Month$ is more strongly significant (a conclusion that matches what we saw in the boxplots earlier). Notice that we cannot run a test for an interaction between the two variables, because we have only one observation per combination of levels.