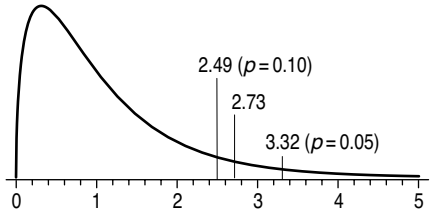
Chapter 13 Solutions

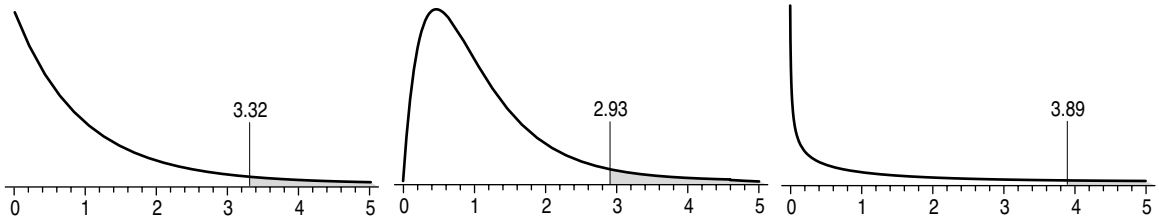
13.1. **(a)** Two-way ANOVA is used when there are two *explanatory* variables. **(b)** Each level of A should occur with all three levels of B. **(c)** The RESIDUAL part of the model represents the error.

13.2. **(a)** It is not necessary that all sample sizes be the same. (The *standard deviations* must all be the same.) **(b)** s_p^2 is found by pooling the sample variances for each SRS. **(c)** The main effects can give useful information even in the presence of an interaction. (This is the "Caution" noted on page 691.)

13.3. A 2×3 ANOVA with 6 observations per cell has $I = 2$, $J = 3$, and $N = 36$. **(a)** The degrees of freedom for interaction are DFAB $= (I-1)(J-1) = 2$ and DFE $= N - IJ = 30$. **(b)** The sketch on the right shows the observed F-value given in part (c) and the bounding critical values from Table E. **(c)** In Table E, we see that $2.49 < F < 3.32$, so $0.05 < P < 0.10$. (Software gives 0.0814.) **(d)** The plot of means would look somewhat parallel because the interaction term is not significantly different from 0 at $\alpha = 0.05$. (Because the P-value is pretty small, the means will not be perfectly parallel.)



13.4. The answers are found in Table E (or using software) with $p = 0.05$. **(a)** We have $I = 3$, $J = 5$ and $N = 45$, so DFA $= 2$ and DFE $= 30$. We would reject H_0 if $F > 3.32$ (software gives 3.3158). **(b)** We have $I = J = 3$ and $N = 27$, so DFAB $= 4$ and DFE $= 18$. We would reject H_0 if $F > 2.93$ (software: 2.9278). **(c)** We have $I = J = 2$ and $N = 204$, so DFAB $= 1$ and DFE $= 200$. We would reject H_0 if $F > 3.89$ (software: 3.8884).

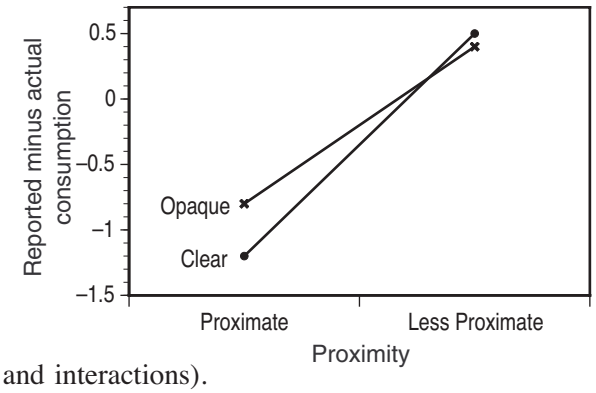


13.5. **(a)** The factors are gender ($I = 2$) and age ($J = 3$). The response variable is the percent of pretend play. The total number of observations is $N = (2)(3)(11) = 66$. **(b)** The factors are time after harvest ($I = 5$) and amount of water ($J = 2$). The response variable is the percent of seed germinating. The total number of observations is $N = 30$ (3 lots of seeds in each of the 10 treatment combinations). **(c)** The factors are media type ($I = 2$) and incubation time ($J = 3$). The response variable is the growth of the virus. The total number of observations is $N = (2)(3)(4) = 24$.

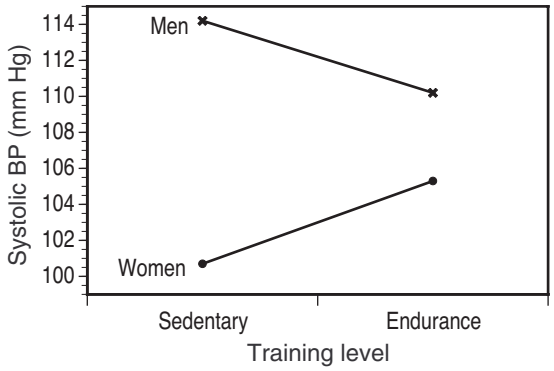
13.6. The table on the right summarizes the degrees of freedom for each source.

		(a)	(b)	(c)
	$I =$	2	5	2
	$J =$	3	2	3
Source	$N =$	66	30	24
A	$I - 1 =$	1	4	1
B	$J - 1 =$	2	1	2
AB	$(I - 1)(J - 1) =$	2	4	2
Error	$N - IJ =$	60	20	18

13.7. (a) The plot suggests a possible interaction because the means are not parallel. (Note that we could have chosen to put dish type on the horizontal axis instead of proximity; either explanatory variable will do.) **(b)** By subjecting the same individual to all four treatments, rather than four individuals to one treatment each, we reduce the within-groups variability (the residual), which makes it easier to detect between-groups variability (the main effects and interactions).



13.8. (a) The plot suggests a gender effect: Men had higher postexercise blood pressure (BP) than women. There also appears to be an interaction: BP was higher for endurance-trained women than for sedentary women (as the researchers had hypothesized), but for men, that pattern was reversed. **(b)** The complete ANOVA table is given in the Minitab output below. The apparent interaction noted in (a) was not significant, but there is a significant effect of gender. **(c)** Subjects with a high pre-exercise BP are likely to have higher postexercise BP, as well. By incorporating both measurements, the researchers can focus of the *change* in BP after exercising, which should be a better measure of the effect of exercise.



Note: *The fact that the interaction was not significant, despite the appearance of the plot, is due to the large variation in individual BPs, indicated by the sizes of the standard errors given in the table. (Observe that these are standard* errors, *not standard* deviations.*) One reason for measuring change in BP—as suggested in (c)—is that we might expect this measurement to have less subject-to-subject variation.*

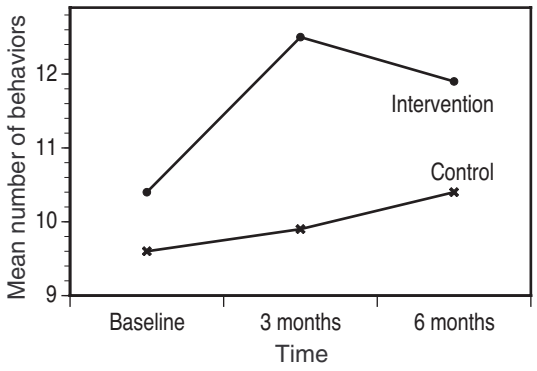
Minitab output

```
Analysis of Variance for BP

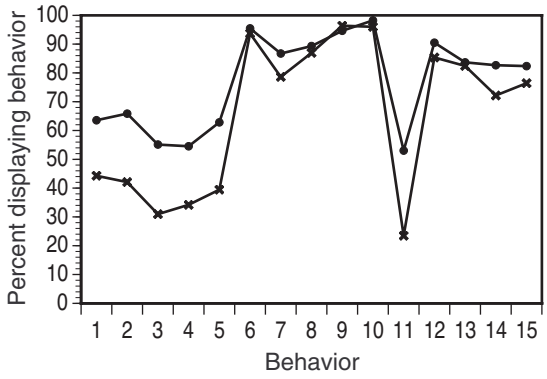
Source             DF         SS         MS       F      P
Gender              1     677.12     677.12    7.65  0.010
Training            1       0.72       0.72    0.01  0.929
Gender*Training     1     147.92     147.92    1.67  0.207
Error              28    2478.00      88.50
Total              31    3303.76
```

13.9. **(a)** The marginal means (as well as the individual cell means) are in the table below. The first two means suggest that the intervention group showed more improvement than the control group. **(b)** Interaction means that the mean number of actions changes differently over time for the two groups. We see this in the plot below because the lines connecting the means are not parallel.

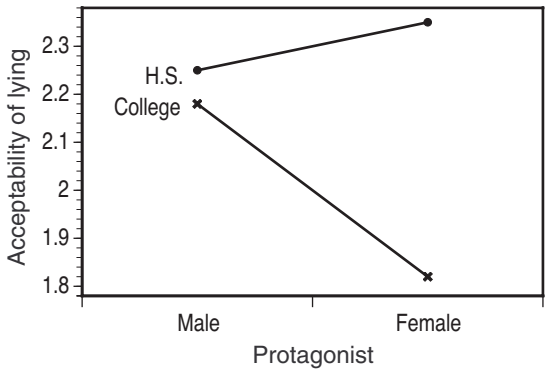
	Time			
Group	Baseline	3 mo.	6 mo.	Mean
Intervention	10.4	12.5	11.9	11.6
Control	9.6	9.9	10.4	9.967
Mean	10.0	11.2	11.15	10.783



13.10. **(a)** The data might be displayed in a variety of ways. Because there are so many numbers (intervention and control groups, at baseline, 3 months, and 6 months), the graph can very easily become overwhelmingly crowded; in order to avoid this, the graph on the right shows the percentage for each group averaged over the three times. Any reasonable graphical display will likely be judged more effective than Table 13.1; for the most part, it is easier to interpret pictures than lists of numbers. **(b)** The behaviors seemed to fall into two categories: Those that both groups did most of the time and those that were less common. The biggest differences between the control and intervention groups are in the latter group, which includes the first five and the 11th behaviors: hide money, hide extra keys, abuse code to alert family, hide extra clothing, asked neighbors to call police, removed weapons. These behaviors should receive special attention in future programs. **(c)** Perhaps the results of this study may be less applicable to smaller communities, or to those which are less diverse.

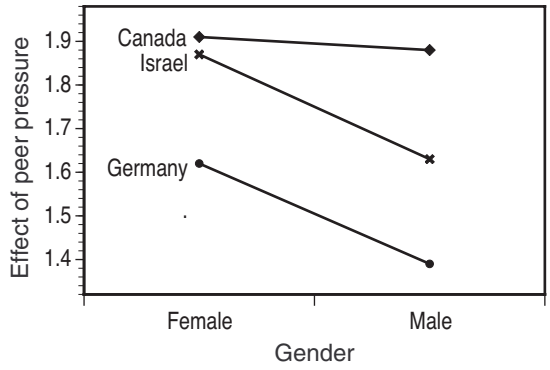


13.11. **(a)** High school and college results are very similar for male protagonists, but there is an apparent interaction: College students appear to be less accepting of lying female protagonists. (Note that we could have chosen to put age on the horizontal axis instead of the gender of the protagonist.)



(b) With $I = J = 2$ and $N = 490$, the degrees of freedom for testing the interaction are DFAB $= (I-1)(J-1) = 1$ and DFE $= N - IJ = 486$. Comparing $F = 3.26$ to the $F(1, 200)$ distribution in Table E, we see that $2.73 < F < 3.89$, so $0.05 < P < 0.10$ (software gives 0.0716). **(c)** The results might generalize to high schools and universities that are similar geographically and demographically, but it would be risky to assume they apply to, for example, East coast high schools or universities.

13.12. **(a)** The significant interaction ($P = 0.016$) is visible in the plot of means: males and females in Canada report similar amounts of peer pressure, while males in Germany and Israel report less peer pressure than females. (Note that we could have chosen to put country on the horizontal axis instead of gender.)



(b) It is not entirely clear how the value placed on achievement relates to fear of being called a nerd or teacher's pet. One view might be that in a culture that values achievement, peers should be less likely to "punish" those who do well. If we take this view, the data would refute the researchers' hypothesis. One could also argue that placing a high value on achievement might make students more competitive, and name-calling might be an expression of that competitiveness. Under this view, the low peer-pressure scores for Germany support the researchers' hypothesis. **(c)** With both responses, we could explore the relationship between achievement and fear of peer pressure—for example, are high-achieving students more or less concerned about peer pressure than low-achieving students?

13.13. We have $I = 3$, $J = 2$ and $N = 30$, so the degrees of freedom are DFA $= 2$, DFB $= 1$, DFAB $= 2$, and DFE $= 24$. This allows us to determine P-values (or to compare to Table E), and we find that there are no significant effects (although B and AB are close):

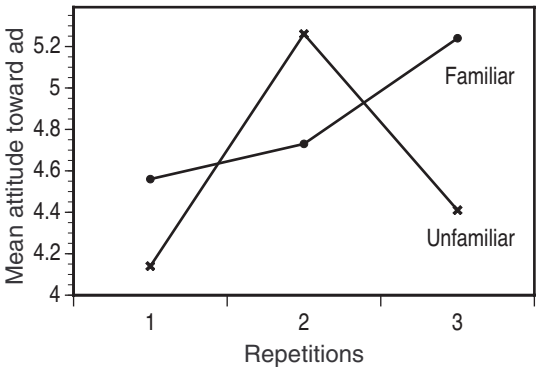
$$F_{\text{A}} = 1.53 \text{ has df 2 and 24, so } P = 0.2369$$

$$F_{\text{B}} = 3.87 \text{ has df 1 and 24, so } P = 0.0608$$

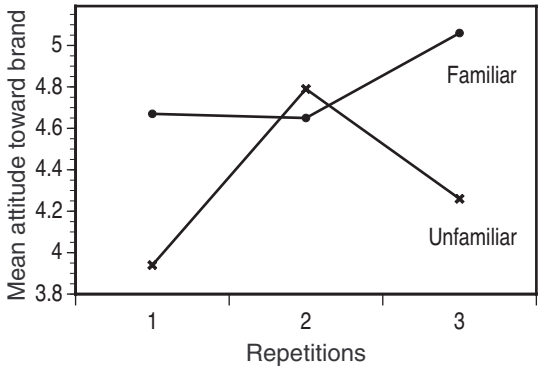
$$F_{\text{AB}} = 2.94 \text{ has df 2 and 24, so } P = 0.0721$$

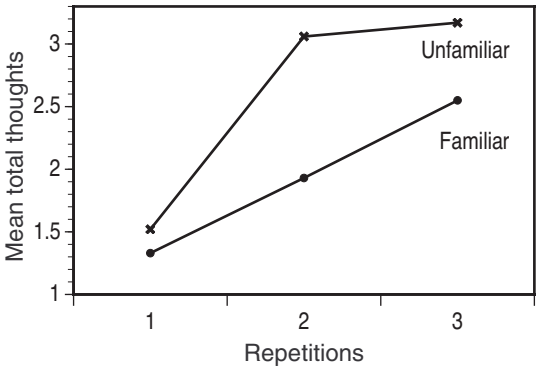
13.14. **(a)** Based on the given P-values, all three effects are significant at $\alpha = 0.05$. (The interaction effect would not be significant at $\alpha = 0.01$.) **(b)** In order to summarize the results, we would need to know the number of levels for each factor (I and J) and the sample sizes in each cell (n_{ij}). We also would want to know the sample cell means $\bar{x}_{ij}$ so that we could interpret the significant main effects and the nature of the interaction.

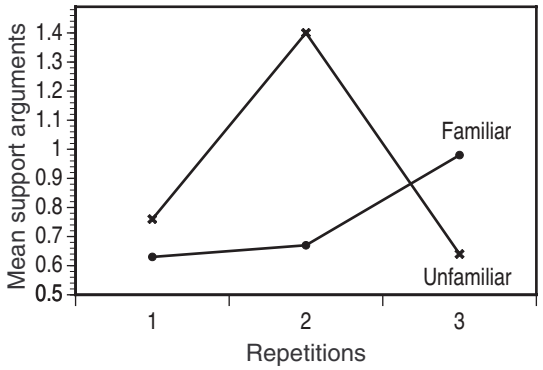
13.15. **(a)** For familiar ads, mean attitude increases with repetition. Familiar ads had the higher mean rating for 1 and 3 repetitions, while unfamiliar ads were rated higher for 2 repetitions. **(b)** The interaction is that unfamiliar does better with 2 repetitions, and worse for 1 and 3. (If there were no interaction, the lines connecting the means would be close to parallel.)

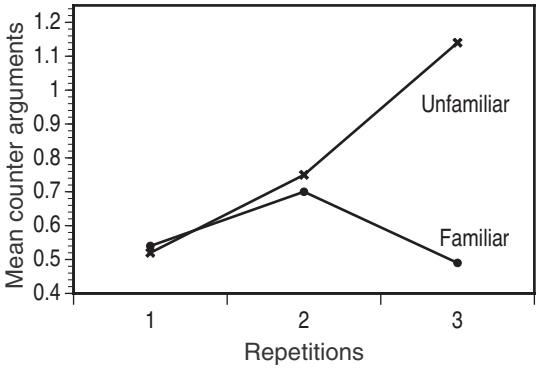


13.16. All four mean plots show evidence of interaction (the lines connecting the means are not parallel). The first and third exhibit a pattern similar to the plot in the previous exercise: For familiar ads, the response variable increases (or decreases only slightly) with increasing repetition, while for unfamiliar ads, the response variable is sharply higher for 2 repetitions. Mean counter arguments displays nearly the opposite pattern: increasing for unfamiliar ads and higher at 2 repetitions for familiar ads. Mean total thoughts increases with repetitions for both familiar and unfamiliar ads.









13.17. Pooling is reasonable, as $1.46/1.16 \doteq 1.26 < 2$. If we assume that all sample sizes were equal, the pooled standard deviation is $s_p \doteq 1.3077$. As long as the sample sizes were approximately equal, we get a similar answer.

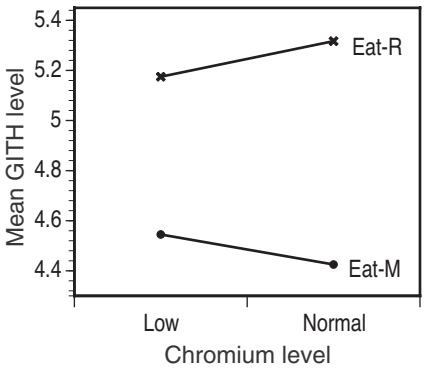
Note: *The published paper from which these numbers were taken only mentions that there were 94 subjects, randomly assigned into groups. Also, the value of s_p is the same regardless of the sample size; that is, we do not need to know the sample sizes in order to compute this. It is simplest to do the computation assuming that all sample sizes were 2; in that case, we simply square the six given standard deviations, find the average, and then take the square root.*

13.18. Pooling is reasonable, as $2.16/1.42 \doteq 1.52 < 2$. If we assume that all sample sizes were equal (see the comment in the previous solution), the pooled standard deviation is $s_p \doteq 1.6767$.

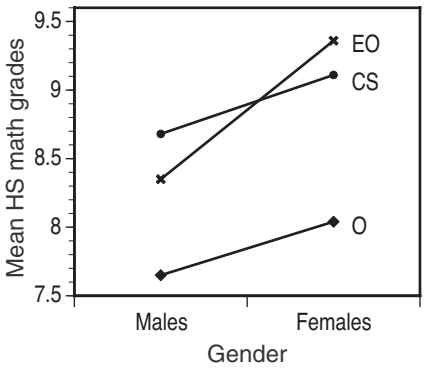
13.19. Questions raised by this background information might include: Are opinions of university employees (and/or West Coast residents) similar to those of other groups? What do experts consider to be "good" ads? Did the content of the news show affect responses to the ads?

13.20. The data cannot be Normally distributed, as we have whole numbers from 1 to 7. However, there can be no extreme outliers on such a scale, and we have averaged the ratings of three ads, so ANOVA would be fairly safe.

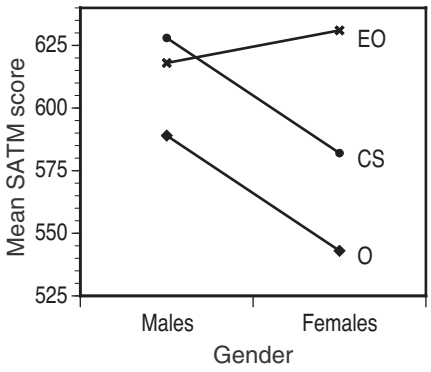
13.21. **(a)** Plot on the right. **(b)** There seems to be a fairly large difference between the means based on how much the rats were allowed to eat, but not very much difference based on the chromium level. There may be an interaction: the NM mean is lower than the LM mean, while the NR mean is higher than the LR mean. **(c)** The marginal means are L: 4.86, N: 4.871, M: 4.485, R: 5.246. For low chromium level (L), R minus M is 0.63; for normal chromium (N), R minus M is 0.892. Mean GITH levels are lower for M than for R; there is not much difference for L *versus* N. The difference between M and R is greater among rats who had normal chromium levels in their diets (N).



13.22. The "Other" category had the lowest mean HS math grades for both genders; this is apparent from the graph (right) as well as from the marginal means (CS: 8.895, EO: 8.855, O: 7.845). Females had higher mean grades; the female marginal mean is $8.83\overline{6}$ compared to $8.22\overline{6}$ for males. The female − male difference is similar for CS and O (about 0.5) but is about twice as big for EO (an interaction).



13.23. The "Other" category had the lowest mean SATM score for both genders; this is apparent from the graph (right) as well as from the marginal means (CS: 605, EO: 624.5, O: 566.) Males had higher mean scores in CS and O, while females are slightly higher in EO; this indicates an interaction. Overall, the marginal means by gender are 611.7 (males) and 585.3 (females).

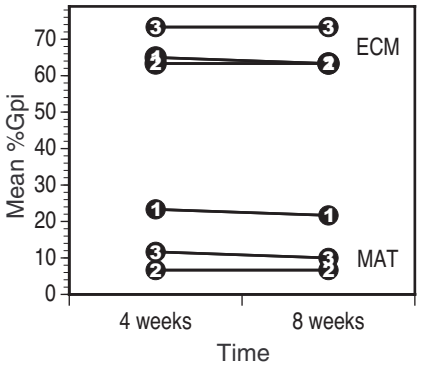


13.24. A study today might include a category for those who declared a major such as Information Technology (which probably did not exist at the time of the initial study). Some variables that might be useful to consider: grade in first programming course, high school physics grades, etc.

13.25. **(a)** For all cells, $n = 3$; the means, standard deviations, and standard errors are in the table below. The ratio of the largest to smallest standard deviations is $8.66/2.88 = 3$, but apart from the first standard deviation, the ratio is 2. **(b)** ECM means are all higher than MAT means. Time and interaction effects are not clearly suggested in the plot (the means are pretty close to parallel, and for a given material, there is little difference between 4 and 8 weeks). **(c)** As the plot suggests, only Material ($F = 251.26$, df 5 and 24, $P < 0.0005$) is significant. For Time, $F = 0.29$, df 1 and 24, $P = 0.595$. For interaction, $F = 0.06$, df 5 and 24, $P = 0.998$. Based on this, we see clearly that the best results come from ECM scaffolds; no additional benefit is obtained from leaving the scaffold in place for eight weeks.

Gpi (%)

	4 weeks			8 weeks		
Material	$\bar{x}$	s	$s_{\bar{x}}$	$\bar{x}$	s	$s_{\bar{x}}$
ECM1	65	8.6603	5	$63.\overline{3}$	2.8868	$1.\overline{6}$
ECM2	$63.\overline{3}$	2.8868	$1.\overline{6}$	$63.\overline{3}$	5.7735	$3.\overline{3}$
ECM3	$73.\overline{3}$	2.8868	$1.\overline{6}$	$73.\overline{3}$	5.7735	$3.\overline{3}$
MAT1	$23.\overline{3}$	2.8868	$1.\overline{6}$	$21.\overline{6}$	5.7735	$3.\overline{3}$
MAT2	$6.\overline{6}$	2.8868	$1.\overline{6}$	$6.\overline{6}$	2.8868	$1.\overline{6}$
MAT3	$11.\overline{6}$	2.8868	$1.\overline{6}$	10	5	2.8868



Minitab output

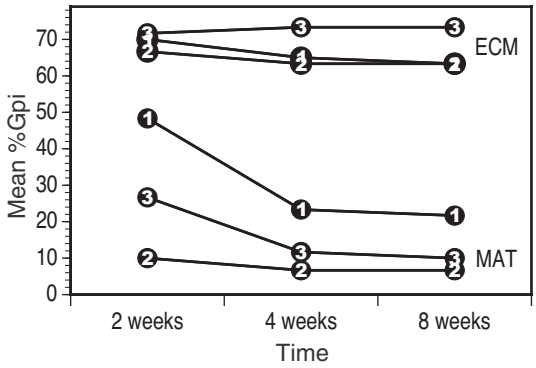
```
Analysis of Variance for Gpi

Source           DF        SS        MS       F      P
Material          5   27045.1    5409.0  251.26  0.000
Weeks             1       6.2       6.2    0.29  0.595
Weeks*Material    5       6.2       1.3    0.06  0.998
Error            24     516.7      21.5
Total            35   27574.3
```

13.26. The means, standard deviations, and standard errors for two weeks are in the table below; as in the previous exercise, $n = 3$ for all cells. Because of the large standard deviation for ECM3 at 2 weeks, the ratio of the largest to smallest standard deviations is now 3.6. The means plot again shows that ECM means are all higher than MAT means. In this plot, there appears to be a time effect (%Gpi decreases over time, at least for MAT1 and MAT3) and an interaction (the amount of decrease varies). ANOVA reveals that all three effects are significant: material ($F = 262.65$, df 5 and 36, $P < 0.0005$), time ($F = 16.93$, df 2 and 36, $P < 0.0005$), and interaction ($F = 3.81$, df 10 and 36, $P = 0.001$). Our conclusions are still the same (ECM scaffolds are better, and for ECM, time does not matter), but the addition of the 2-week data adds a cautionary note: Not only are the inert-material scaffolds MAT1, MAT2, and MAT3 less effective than the ECM scaffolds, but they seem to do worse over time.

Gpi (%) at 2 weeks

Material	$\bar{x}$	s	$s_{\bar{x}}$
ECM1	70	5	2.8868
ECM2	$66.\overline{6}$	7.6376	4.4096
ECM3	$71.\overline{6}$	10.4083	6.0093
MAT1	$48.\overline{3}$	2.8868	$1.\overline{6}$
MAT2	10	5	2.8868
MAT3	$26.\overline{6}$	2.8868	$1.\overline{6}$



Minitab output

```
Analysis of Variance for Gpi

Source           DF        SS        MS       F      P
Material          5   35870.8    7174.2  262.65  0.000
Weeks             2     925.0     462.5   16.93  0.000
Material*Weeks   10    1041.7     104.2    3.81  0.001
Error            36     983.3      27.3
Total            53   38820.8
```

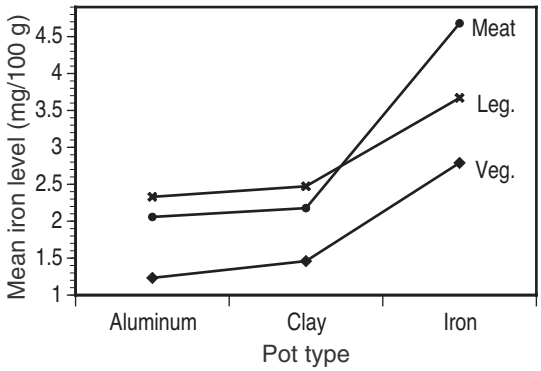
13.27. For each time period, there is a significant difference among materials (Minitab output below). The one-way ANOVA for the 4-week time period also appears in the solution to Exercise 12.51. With df = 12, 15 comparisons, and $\alpha = 0.05$, the Bonferroni critical value is $t^{**} = 3.6489$. The table on the right gives the pooled standard deviations s_p, the standard errors of each difference $\text{SE}_D = s_p\sqrt{1/3 + 1/3}$, and the "minimum significant difference" $\text{MSD} = t^{**}\text{SE}_D$ (two means are significantly different if they differ by at least this amount). For a given time period, the ECM means are not significantly different from one another. MAT1 is significantly larger than MAT2 for all time periods, and is also greater than MAT3 at 2 weeks. The only ECM/MAT difference that is *not* significant is ECM2 and MAT1 at 2 weeks.

Time period	s_p	SE_D	MSD
2 weeks	6.2361	5.0918	18.5793
4 weeks	4.4096	3.6004	13.1375
8 weeks	4.8591	3.9675	14.4768

Minitab output

```
- - - - - - - - - - - - - - - - - -  Two weeks  - - - - - - - - - - - - - - - - - -
Analysis of Variance on Gpi2
Source     DF        SS        MS        F        p
Material    5    9861.1    1972.2    50.71    0.000
Error      12     466.7      38.9
Total      17   10327.8
- - - - - - - - - - - - - - - - - -  Four weeks  - - - - - - - - - - - - - - - - - -
Analysis of Variance on Gpi4
Source     DF        SS        MS        F        p
Material    5   13411.1    2682.2   137.94    0.000
Error      12     233.3      19.4
Total      17   13644.4
- - - - - - - - - - - - - - - - - -  Eight weeks  - - - - - - - - - - - - - - - - - -
Analysis of Variance on Gpi8
Source     DF        SS        MS        F        p
Material    5   13640.3    2728.1   115.54    0.000
Error      12     283.3      23.6
Total      17   13923.6
```

13.28. **(a)** The sample size is $n = 4$ for each pot/food combination; means and standard deviations are given in the table below. The largest-to-smallest ratio is $0.63/0.07 \doteq 8.8$, which is well above our guideline for pooling. **(b)** The iron levels differed among the three food types, and for all food types, aluminum and clay pots produced similar iron levels, while iron pots resulted in much higher iron levels. There is also evidence of an interaction: Iron levels in iron pots rose much more for meat than for legumes or vegetables. **(c)** The ANOVA table (below) shows that all three effects are quite significant.



Iron (mg per 100 g)

	Meat		Legumes		Vegetables	
Pot type	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s
Aluminum	2.0575	0.2520	2.3300	0.1111	1.2325	0.2313
Clay	2.1775	0.6213	2.4725	0.0714	1.4600	0.4601
Iron	4.6800	0.6283	3.6700	0.1726	2.7900	0.2399

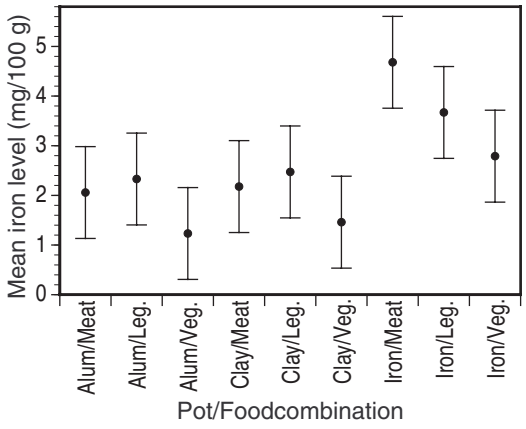
Minitab output

```
Analysis of Variance for Iron

Source      DF        SS        MS       F      P
Pot          2   24.8940   12.4470   92.26  0.000
Food         2    9.2969    4.6484   34.46  0.000
Pot*Food     4    2.6404    0.6601    4.89  0.004
Error       27    3.6425    0.1349
Total       35   40.4738
```

13.29. Yes; the iron-pot means are the highest, and the F statistic for testing the effect of the pot type is very large. (In this case, the interaction does not weaken any evidence that iron-pot foods contain more iron; it only suggests that while iron pots increase iron levels in all foods, the effect is strongest for meats.)

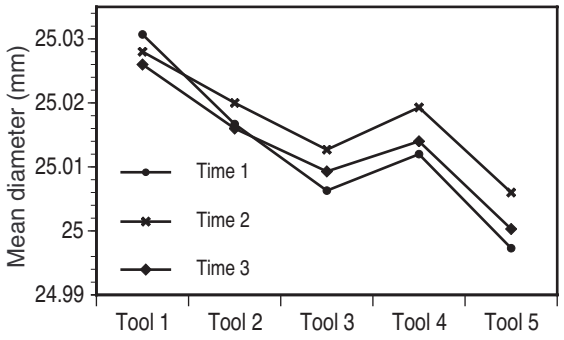
13.30. The ANOVA table (below) shows significant evidence that at least one group mean is different. With df $= 27, 36$ comparisons, and $\alpha = 0.05$, the Bonferroni critical value is $t^{**} = 3.5629$. The pooled standard deviation is $s_p \doteq 0.3673$, the standard error of each difference $\text{SE}_D = s_p\sqrt{1/4 + 1/4} \doteq 0.2597$, so two means are significantly different if they differ by $t^{**}\text{SE}_D \doteq 0.9253$. The "error bars" in the plot on the right are drawn with this length (above and below each mean), so two means are significantly different if the "dot" for one mean does not fall within the other mean's error bars. For example, we find that iron/meat is significantly larger than everything else, and iron/legumes is significantly different from everything except iron/vegetable. These conclusions are consistent with the results of the two-way ANOVA.



Minitab output

```
Analysis of Variance on Iron
Source     DF        SS        MS        F        p
Potfood     8    36.831     4.604    34.13    0.000
Error      27     3.643     0.135
Total      35    40.474
```

13.31. **(a)** For all tool/time combinations, $n = 3$. Means and standard deviations are in the table below. Note that five cells had no variability ($s = 0$). **(b)** Plot on the right. Except for tool 1, mean diameter is highest at time 2. Tool 1 had the highest mean diameters, followed by tool 2, tool 4, tool 3, and tool 5. **(c)** Minitab output below; all F statistics are highly significant. **(d)** There is strong evidence of a difference in mean diameter among the tools and among the times. There is also an interaction (specifically, tool 1's mean diameters changed differently over time compared to the other tools).



Minitab output

```
Analysis of Variance for Diameter

Source       DF         SS         MS       F      P
Tool          4 0.00359720 0.00089930  412.94  0.000
Time          2 0.00018991 0.00009496   43.60  0.000
Tool*Time     8 0.00013320 0.00001665    7.65  0.000
Error        30 0.00006533 0.00000218
Total        44 0.00398564
```

Diameter (mm)

	Time 1 (8:00AM)		Time 2 (11:00AM)		Time 3 (3:00PM)	
Tool	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s
1	25.0307	0.001155	25.0280	0	25.0260	0
2	25.0167	0.001155	25.0200	0.002000	25.0160	0
3	25.0063	0.001528	25.0127	0.001155	25.0093	0.001155
4	25.0120	0	25.0193	0.001155	25.0140	0.004000
5	24.9973	0.001155	25.0060	0	25.0003	0.001528

13.32. All means and standard deviations will change by a factor of 0.04; the plot is identical to that in Exercise 13.31, except that the vertical scale is different. All SS and MS values change by a factor of $0.04^2 = 0.0016$, but the F- (and P-) values are the same.

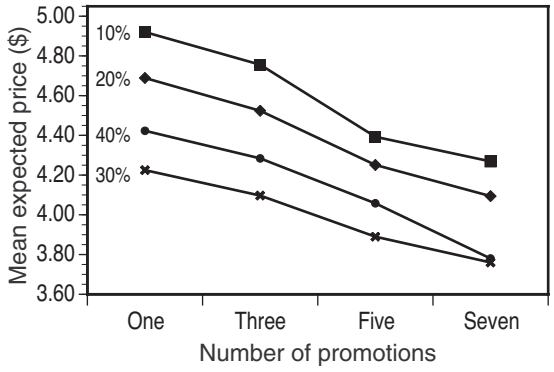
Minitab output

```
Two-way ANOVA: diameter versus tool, time

Source       DF         SS         MS       F      P
tool          4  0.0000058  0.0000014  412.94  0.000
time          2  0.0000003  0.0000002   43.60  0.000
Interaction   8  0.0000002  0.0000000    7.65  0.000
Error        30  0.0000001  0.0000000
Total        44  0.0000064

S = 0.00005903   R-Sq = 98.36%   R-Sq(adj) = 97.60%
```

13.33. **(a)** Table below, plot on the right. The mean expected price decreases as the number of promotions increases and also as the percent discount increases up to 30%. The expected price is higher for 40% than for 30%. **(b)** Minitab output on the next page. Both main effects are significant, but there was no significant interaction. **(c)** There is strong evidence of a difference in mean expected price based on the number of promotions and the percent discount. Specifically, the two effects noted in (a) are significant: More promotions and higher discounts (up to a point) decrease the expected price. There is no evidence of an interaction.



Expected price ($)

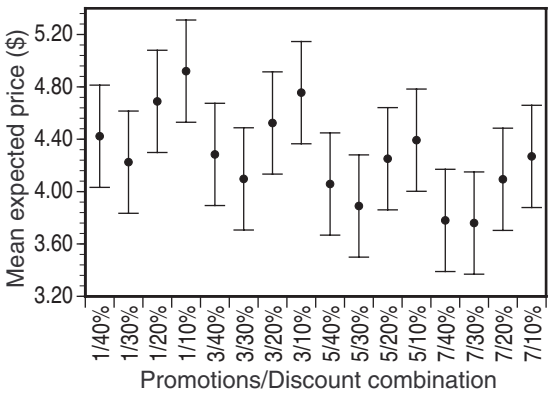
	40% discount		30% discount		20% discount		10% discount	
Promotions	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s	$\bar{x}$	s
1	4.423	0.1848	4.225	0.3856	4.689	0.2331	4.920	0.1520
3	4.284	0.2040	4.097	0.2346	4.524	0.2707	4.756	0.2429
5	4.058	0.1760	3.890	0.1629	4.251	0.2648	4.393	0.2685
7	3.780	0.2144	3.760	0.2618	4.094	0.2407	4.269	0.2699

Minitab output

```
Analysis of Variance for ExpPrice

Source             DF         SS         MS       F      P
Promos              3     8.3605     2.7868   47.73  0.000
Discount            3     8.3069     2.7690   47.42  0.000
Promos*Discount     9     0.2306     0.0256    0.44  0.912
Error             144     8.4087     0.0584
Total             159    25.3067
```

13.34. ANOVA gives $F = 19.29$ (df 15 and 144), which has $P < 0.0005$—there is significant evidence that some means are different. With df = 144, 120 comparisons, and $\alpha = 0.05$, the Bonferroni critical value is $t^{**} = 3.6135$. The pooled standard deviation is $s_p \doteq 0.2416$, and the standard error of each difference $\text{SE}_D = s_p\sqrt{1/10 + 1/10} \doteq 0.1081$, so two means are significantly different if they differ by $t^{**}\text{SE}_D \doteq 0.3905$. The "error bars" in the plot on the right are drawn with this length (above and below each mean), so two means are significantly different if the "dot" for one mean does not fall within the other mean's error bars. For example, we find that the 1 promotion/10% discount mean differs from all means except 1/20% and 3/10%.



Minitab output

```
Analysis of Variance on ExpPrice
Source     DF        SS        MS        F        p
PromDisc   15   16.8980    1.1265    19.29    0.000
Error     144    8.4087    0.0584
Total     159   25.3067
```

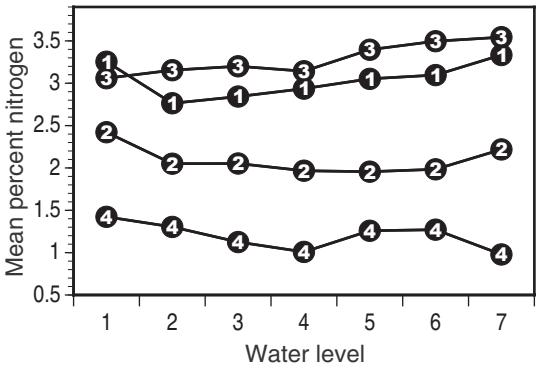
13.35. (a) All three F values have df 1 and 945, and the P-values are < 0.001, < 0.001, and 0.1477. Gender and handedness both have significant effects on mean lifetime, but there is no significant interaction. **(b)** Women live about 6 years longer than men (on the average), while right-handed people average 9 more years of life than left-handed people. "There is no interaction" means that handedness affects both genders in the same way, and vice versa.

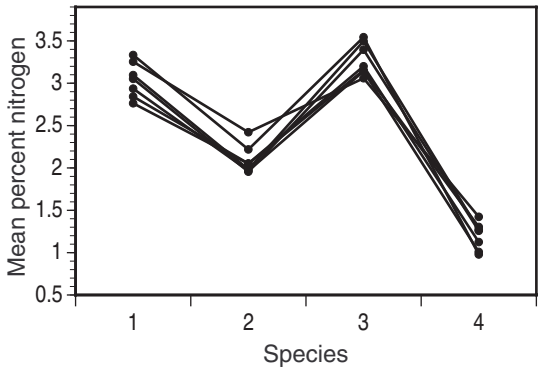
13.36. (a) $F_{\text{series}} = 7.02$ with df 3 and 61; this has $P = 0.0004$. $F_{\text{holder}} = 1.96$ with df 1 and 61; this has $P = 0.1665$. $F_{\text{interaction}} = 1.24$ with df 3 and 61; this has $P = 0.3026$. Only the series had a significant effect; the presence or absence of a holder and series/holder interaction did not significantly affect the mean radon reading. **(b)** Because the ANOVA indicates that these means are significantly different, we conclude that detectors produced in different production runs give different readings for the same radon level. This inconsistency may indicate poor quality control in production.

Note: *In the initial printing of the text, the total sample size (N = 69) was not given, without which we cannot determine the denominator degrees of freedom for part (a).*

13.37. (a) & **(b)** The table below lists the means and standard deviations (the latter in parentheses). The two plots below suggest that plant 1 and plant 3 have the highest nitrogen content, plant 2 is in the middle, and plant 4 is the lowest. (In the second plot, the points are so crowded together that no attempt was made to differentiate among the different water levels.) There is no consistent effect of water level on nitrogen content. Standard deviations range from 0.0666 to 0.3437, for a ratio of 5.16—larger than we like. **(c)** Minitab output below. Both main effects and the interaction are highly significant.

	Percent nitrogen						
Species	50mm	150mm	250mm	350mm	450mm	550mm	650mm
1	3.2543 (0.2287)	2.7636 (0.0666)	2.8429 (0.2333)	2.9362 (0.0709)	3.0519 (0.0909)	3.0963 (0.0815)	3.3334 (0.2482)
2	2.4216 (0.1654)	2.0502 (0.1454)	2.0524 (0.1481)	1.9673 (0.2203)	1.9560 (0.1571)	1.9839 (0.2895)	2.2184 (0.1238)
3	3.0589 (0.1525)	3.1541 (0.3324)	3.2003 (0.2341)	3.1419 (0.2965)	3.3956 (0.2533)	3.4961 (0.3437)	3.5437 (0.3116)
4	1.4230 (0.1738)	1.3037 (0.2661)	1.1253 (0.1230)	1.0087 (0.1310)	1.2584 (0.2489)	1.2712 (0.0795)	0.9788 (0.2090)



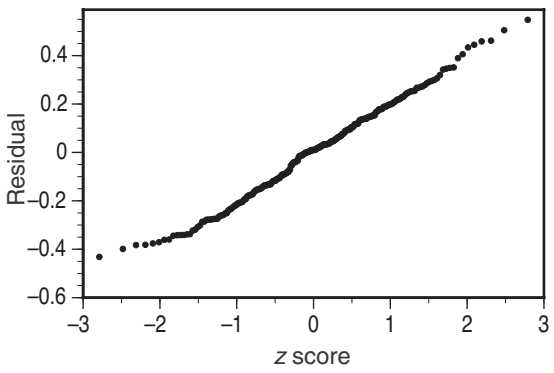


Minitab output

```
Analysis of Variance for Pctnit

Source           DF         SS         MS       F      P
Species           3   172.3916    57.4639 1301.32  0.000
Water             6     2.5866     0.4311    9.76  0.000
Species*Water    18     4.7446     0.2636    5.97  0.000
Error           224     9.8914     0.0442
Total           251   189.6143
```

13.38. The residuals appear to be reasonably Normal, with no apparent outliers and no clear patterns.



13.39. For each water level, there is highly significant evidence of variation in nitrogen level among plant species (Minitab output below). For each water level, we have df $= 32$, 6 comparisons, and $\alpha = 0.05$, so the Bonferroni critical value is $t^{**} = 2.8123$. (If we take into account that there are 7 water levels, so that overall we are performing $6 \times 7 = 42$ comparisons, we should take $t^{**} = 3.5579$.) The table on the right gives the pooled standard deviations s_p, the standard errors of each difference $\text{SE}_D = s_p\sqrt{1/9 + 1/9}$, and the "minimum significant difference" $\text{MSD} = t^{**}\text{SE}_D$ (two means are significantly different if they differ by at least this amount). MSD1 uses $t^{**} = 2.8123$, and MSD2 uses $t^{**} = 3.5579$. As it happens, for either choice of MSD, the only *non*significant differences are between species 1 and 3 for water levels 1, 4, and 7. (These are the three closest pairs of points in the plot from the solution to Exercise 13.37.) Therefore, for every water level, species 4 has the lowest nitrogen level and species 2 is next. For water levels 1, 4, and 7, species 1 and 3 are statistically tied for the highest level; for the other levels, species 3 is the highest, with species 1 coming in second.

Water level	s_p	SE_D	MSD1	MSD2
1	0.1824	0.0860	0.2418	0.3059
2	0.2274	0.1072	0.3015	0.3814
3	0.1912	0.0902	0.2535	0.3208
4	0.1991	0.0939	0.2640	0.3340
5	0.1994	0.0940	0.2643	0.3344
6	0.2318	0.1093	0.3073	0.3887
7	0.2333	0.1100	0.3093	0.3913

Minitab output

Water level 1

Source	DF	SS	MS	F	p
Species	3	18.3711	6.1237	184.05	0.000
Error	32	1.0647	0.0333		
Total	35	19.4358			

Water level 2

Source	DF	SS	MS	F	p
Species	3	17.9836	5.9945	115.93	0.000
Error	32	1.6546	0.0517		
Total	35	19.6382			

Water level 3

Source	DF	SS	MS	F	p
Species	3	22.9171	7.6390	208.87	0.000
Error	32	1.1704	0.0366		
Total	35	24.0875			

Water level 4

Source	DF	SS	MS	F	p
Species	3	25.9780	8.6593	218.37	0.000
Error	32	1.2689	0.0397		
Total	35	27.2469			

Water level 5

Source	DF	SS	MS	F	p
Species	3	26.2388	8.7463	220.01	0.000
Error	32	1.2721	0.0398		
Total	35	27.5109			

Water level 6

Source	DF	SS	MS	F	p
Species	3	28.0648	9.3549	174.14	0.000
Error	32	1.7191	0.0537		
Total	35	29.7838			

Water level 7

Source	DF	SS	MS	F	p
Species	3	37.5829	12.5276	230.17	0.000
Error	32	1.7417	0.0544		
Total	35	39.3246			

13.40. The F statistics for all four ANOVAs are significant, and all four regressions are significant as well. However, the regressions all have low R^2 (varying from 6.4% to 27.3%), and plots indicate that a straight line is not really appropriate except perhaps for plant 3 (which had the highest R^2 value).

Minitab output

```
- - - - - - - - - - - - - - -   ANOVA: Plant species 1   - - - - - - - - - - - - - - -
Source     DF        SS        MS        F        p
Water       6    2.3527    0.3921    14.25    0.000
Error      56    1.5413    0.0275
Total      62    3.8940
- - - - - - - - - - - - - - -   ANOVA: Plant species 2   - - - - - - - - - - - - - - -
Source     DF        SS        MS        F        p
Water       6    1.5626    0.2604     7.51    0.000
Error      56    1.9420    0.0347
Total      62    3.5046
- - - - - - - - - - - - - - -   ANOVA: Plant species 3   - - - - - - - - - - - - - - -
Source     DF        SS        MS        F        p
Water       6    1.9764    0.3294     4.15    0.002
Error      56    4.4464    0.0794
Total      62    6.4228
- - - - - - - - - - - - - - -   ANOVA: Plant species 4   - - - - - - - - - - - - - - -
Source     DF        SS        MS        F        p
Water       6    1.4396    0.2399     6.85    0.000
Error      56    1.9618    0.0350
Total      62    3.4013
- - - - - - - - - - - - - - -   Regression: Plant species 1   - - - - - - - - - - - - - - -
The regression equation is plant1 = 2.88 + 0.0397 Water

Predictor       Coef       Stdev    t-ratio        p
Constant     2.88097     0.06745      42.71    0.000
Water        0.03971     0.01508       2.63    0.011

s = 0.2394      R-sq = 10.2%     R-sq(adj) = 8.7%
- - - - - - - - - - - - - - -   Regression: Plant species 2   - - - - - - - - - - - - - - -
The regression equation is plant2 = 2.21 - 0.0299 Water

Predictor       Coef       Stdev    t-ratio        p
Constant     2.21262     0.06531      33.88    0.000
Water       -0.02994     0.01460      -2.05    0.045

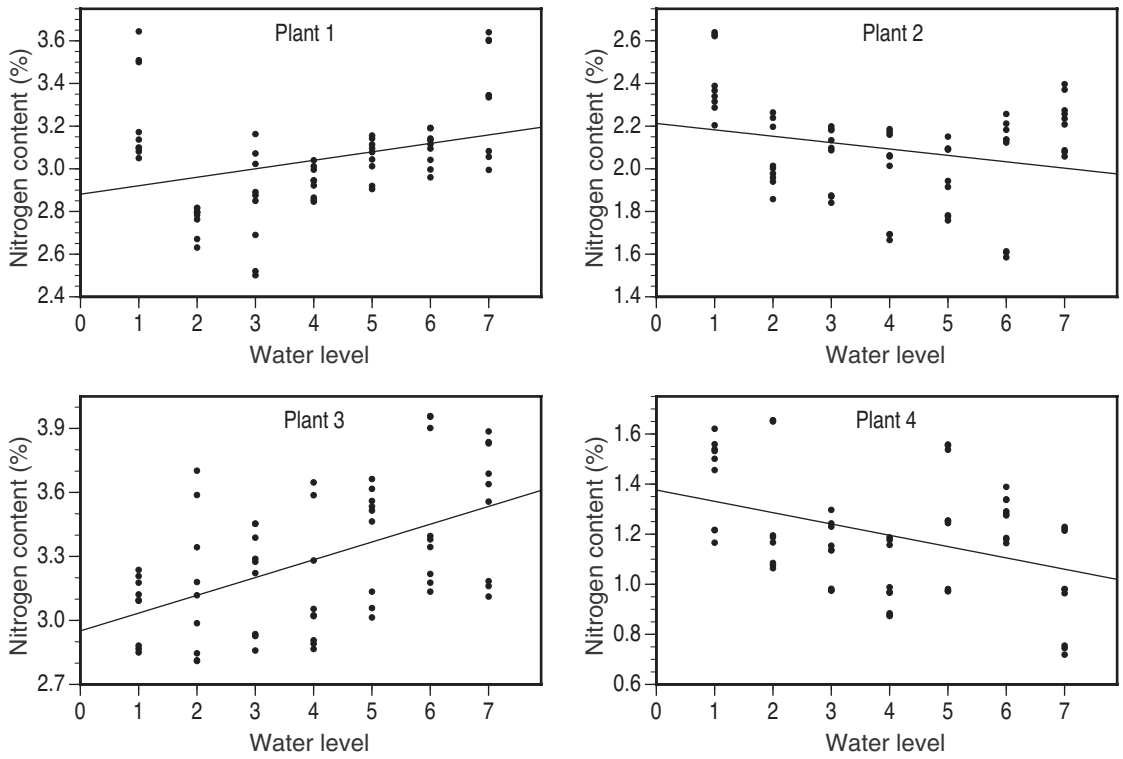
s = 0.2318      R-sq = 6.4%      R-sq(adj) = 4.9%
- - - - - - - - - - - - - - -   Regression: Plant species 3   - - - - - - - - - - - - - - -
The regression equation is plant3 = 2.95 + 0.0833 Water

Predictor       Coef       Stdev    t-ratio        p
Constant     2.95100     0.07797      37.85    0.000
Water        0.08334     0.01743       4.78    0.000

s = 0.2768      R-sq = 27.3%     R-sq(adj) = 26.1%
- - - - - - - - - - - - - - -   Regression: Plant species 4   - - - - - - - - - - - - - - -
The regression equation is plant4 = 1.38 - 0.0452 Water

Predictor       Coef       Stdev    t-ratio        p
Constant     1.37622     0.06129      22.45    0.000
Water       -0.04516     0.01371      -3.29    0.002

s = 0.2176      R-sq = 15.1%     R-sq(adj) = 13.7%
```



13.41. (a) & **(b)** The tables below list the means and standard deviations (the latter in parentheses). The means plots show that biomass (both fresh and dry) increases with water level for all plants. Generally, plants 1 and 2 have higher biomass for each water level, while plants 3 and 4 are lower. Standard deviation ratios are quite high for both fresh and dry biomass: $108.01/6.79 \doteq 15.9$ and $35.76/3.12 \doteq 11.5$. **(c)** Minitab output below. For both fresh and dry biomass, main effects and the interaction are significant. (The interaction for fresh biomass has $P = 0.04$; other P-values are smaller.)

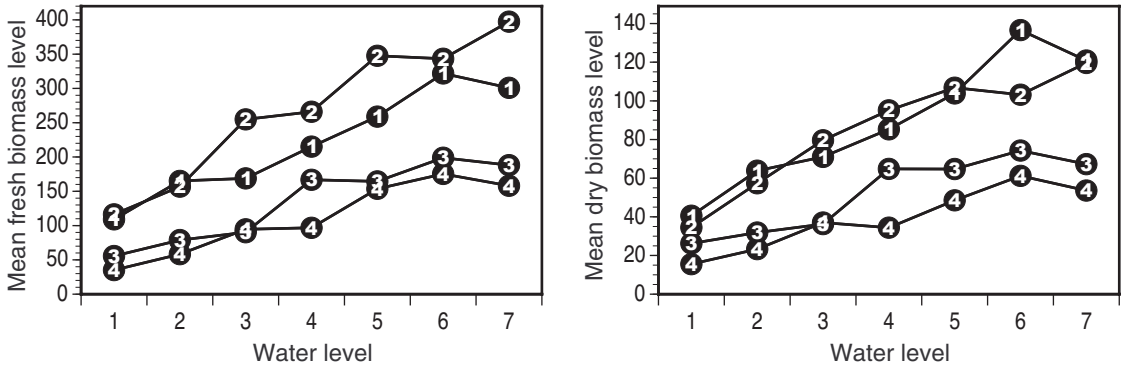
Minitab output

Fresh biomass

Source	DF	SS	MS	F	P
Species	3	458295	152765	81.45	0.000
Water	6	491948	81991	43.71	0.000
Species*Water	18	60334	3352	1.79	0.040
Error	84	157551	1876		
Total	111	1168129			

Dry biomass

Source	DF	SS	MS	F	P
Species	3	50523.8	16841.3	79.93	0.000
Water	6	56623.6	9437.3	44.79	0.000
Species*Water	18	8418.8	467.7	2.22	0.008
Error	84	17698.4	210.7		
Total	111	133264.6			



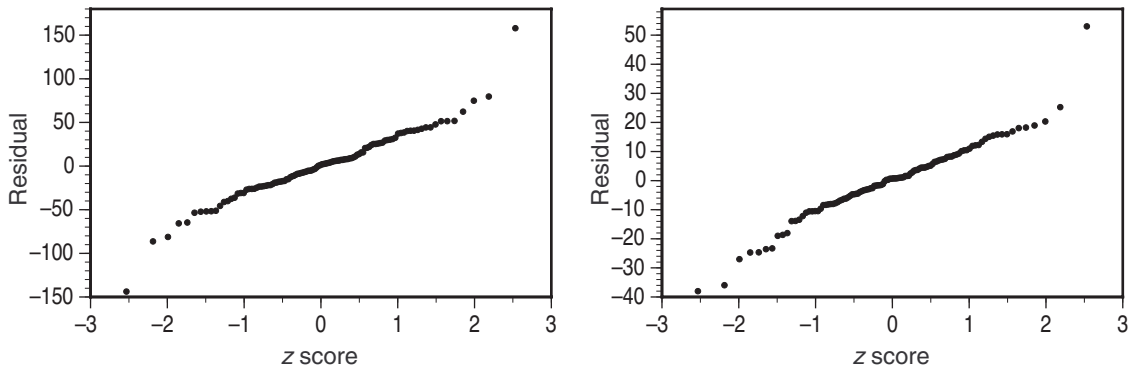
Fresh biomass

Species	50mm	150mm	250mm	350mm	450mm	550mm	650mm
1	109.095	165.138	168.825	215.133	258.900	321.875	300.880
	(20.949)	(29.084)	(18.866)	(42.687)	(45.292)	(46.727)	(29.896)
2	116.398	156.750	254.875	265.995	347.628	343.263	397.365
	(29.250)	(46.922)	(13.944)	(59.686)	(54.416)	(98.553)	(108.011)
3	55.600	78.858	90.300	166.785	164.425	198.910	188.138
	(13.197)	(29.458)	(28.280)	(41.079)	(18.646)	(33.358)	(18.070)
4	35.128	58.325	94.543	96.740	153.648	175.360	158.048
	(11.626)	(6.789)	(13.932)	(24.477)	(22.028)	(32.873)	(70.105)

Dry biomass

Species	50mm	150mm	250mm	350mm	450mm	550mm	650mm
1	40.565	63.863	71.003	85.280	103.850	136.615	120.860
	(5.581)	(7.508)	(6.032)	(10.868)	(15.715)	(16.203)	(17.137)
2	34.495	57.365	79.603	95.098	106.813	103.180	119.625
	(11.612)	(6.149)	(13.094)	(25.198)	(18.347)	(25.606)	(35.764)
3	26.245	31.865	36.238	64.800	64.740	74.285	67.258
	(6.430)	(11.322)	(11.268)	(9.010)	(3.122)	(12.277)	(7.076)
4	15.530	23.290	37.050	34.390	48.538	61.195	53.600
	(4.887)	(3.329)	(5.194)	(11.667)	(5.658)	(12.084)	(25.290)

13.42. Both sets of residuals have a high outlier (observation #53); observation #52 is a low outlier for fresh biomass. The other residuals look reasonably Normal.



13.43. For each water level, there is highly significant evidence of variation in biomass level (both fresh and dry) among plant species (Minitab output below). For each water level, we have df = 12, 6 comparisons, and $\alpha = 0.05$, so the Bonferroni critical value is $t^{**} = 3.1527$. (If we take into account that there are 7 water levels, so that overall we are performing $6 \times 7 = 42$ comparisons, we should take $t^{**} = 4.2192$.) The table below gives the pooled standard deviations s_p, the standard errors of each difference $\text{SE}_D = s_p\sqrt{1/4 + 1/4}$, and the "minimum significant difference" $\text{MSD} = t^{**}\text{SE}_D$ (two means are significantly different if they differ by at least this amount). MSD1 uses $t^{**} = 3.1527$, and MSD2 uses $t^{**} = 4.2192$. Rather than give a full listing of which differences are significant, we note that plants 3 and 4 are *not* significantly different, nor are 1 and 3 (except for one or two water levels). All other plant combinations are significantly different for at least three water levels. For fresh biomass, plants 2 and 4 are different for *all* levels, and for dry biomass, 1 and 4 differ for all levels.

	Fresh biomass				Dry biomass			
Water level	s_p	SE_D	MSD1	MSD2	s_p	SE_D	MSD1	MSD2
1	20.0236	14.1588	44.6382	50.3764	7.6028	5.3760	16.9487	19.1274
2	31.4699	22.2526	70.1552	79.1735	7.6395	5.4019	17.0305	19.2197
3	19.6482	13.8934	43.8012	49.4318	9.5103	6.7248	21.2010	23.9263
4	43.7929	30.9663	97.6265	110.1762	15.5751	11.0133	34.7213	39.1846
5	38.2275	27.0310	85.2197	96.1746	12.5034	8.8412	27.8734	31.4565
6	59.3497	41.9666	132.3068	149.3147	17.4280	12.3235	38.8518	43.8462
7	66.7111	47.1719	148.7174	167.8348	23.7824	16.8167	53.0176	59.8329

Minitab output

Fresh biomass — Water level 1

Source	DF	SS	MS	F	p
Species	3	19107	6369	15.88	0.000
Error	12	4811	401		
Total	15	23918			

Fresh biomass — Water level 2

Source	DF	SS	MS	F	p
Species	3	35100	11700	11.81	0.001
Error	12	11884	990		
Total	15	46984			

Fresh biomass — Water level 3

Source	DF	SS	MS	F	p
Species	3	71898	23966	62.08	0.000
Error	12	4633	386		
Total	15	76531			

Fresh biomass — Water level 4

Source	DF	SS	MS	F	p
Species	3	62337	20779	10.83	0.001
Error	12	23014	1918		
Total	15	85351			

Fresh biomass — Water level 5

Source	DF	SS	MS	F	p
Species	3	99184	33061	22.62	0.000
Error	12	17536	1461		
Total	15	116720			

Fresh biomass — Water level 6

Source	DF	SS	MS	F	p
Species	3	86628	28876	8.20	0.003
Error	12	42269	3522		
Total	15	128897			

Fresh biomass — Water level 7

Source	DF	SS	MS	F	p
Species	3	144376	48125	10.81	0.001
Error	12	53404	4450		
Total	15	197780			

Dry biomass — Water level 1

Source	DF	SS	MS	F	p
Species	3	1411.2	470.4	8.14	0.003
Error	12	693.6	57.8		
Total	15	2104.8			

Dry biomass — Water level 2

Source	DF	SS	MS	F	p
Species	3	4597.1	1532.4	26.26	0.000
Error	12	700.3	58.4		
Total	15	5297.4			

Dry biomass — Water level 3

Source	DF	SS	MS	F	p
Species	3	6127.2	2042.4	22.58	0.000
Error	12	1085.3	90.4		
Total	15	7212.6			

Dry biomass — Water level 4

Source	DF	SS	MS	F	p
Species	3	8634	2878	11.86	0.001
Error	12	2911	243		
Total	15	11545			

Dry biomass — Water level 5

Source	DF	SS	MS	F	p
Species	3	10026	3342	21.38	0.000
Error	12	1876	156		
Total	15	11902			

Dry biomass — Water level 6

Source	DF	SS	MS	F	p
Species	3	13460	4487	14.77	0.000
Error	12	3645	304		
Total	15	17105			

Dry biomass — Water level 7

Source	DF	SS	MS	F	p
Species	3	14687	4896	8.66	0.002
Error	12	6787	566		
Total	15	21474			

13.44. The F statistics for all eight ANOVAs are significant, and all eight regressions are significant as well. Unlike the nitrogen level (Exercises 13.37 through 13.40), all of these regressions have reasonably large values of R^2, and the scatterplots suggest that a straight line is an appropriate model for the relationship.

Minitab output

ANOVA: Fresh biomass, plant species 1

Source	DF	SS	MS	F	p
Water	6	145543	24257	19.76	0.000
Error	21	25774	1227		
Total	27	171317			

ANOVA: Fresh biomass, plant species 2

Source	DF	SS	MS	F	p
Water	6	257083	42847	9.63	0.000
Error	21	93463	4451		
Total	27	350546			

ANOVA: Fresh biomass, plant species 3

Source	DF	SS	MS	F	p
Water	6	80952	13492	17.77	0.000
Error	21	15948	759		
Total	27	96901			

ANOVA: Fresh biomass, plant species 4

Source	DF	SS	MS	F	p
Water	6	68704	11451	10.75	0.000
Error	21	22365	1065		
Total	27	91070			

ANOVA: Dry biomass, plant species 1

Source	DF	SS	MS	F	p
Water	6	27273	4545	30.44	0.000
Error	21	3136	149		
Total	27	30408			

ANOVA: Dry biomass, plant species 2

Source	DF	SS	MS	F	p
Water	6	21802	3634	7.83	0.000
Error	21	9751	464		
Total	27	31553			

ANOVA: Dry biomass, plant species 3

Source	DF	SS	MS	F	p
Water	6	9489.9	1581.6	18.82	0.000
Error	21	1764.6	84.0		
Total	27	11254.5			

ANOVA: Dry biomass, plant species 4

Source	DF	SS	MS	F	p
Water	6	6478	1080	7.44	0.000
Error	21	3047	145		
Total	27	9525			

Regression: Fresh biomass, plant species 1

The regression equation is plant1 = 80.1 + 35.0 Water

Predictor	Coef	Stdev	t-ratio	p
Constant	80.13	15.38	5.21	0.000
Water	34.961	3.438	10.17	0.000

s = 36.39 R-sq = 79.9% R-sq(adj) = 79.1%

```
 - - - - - - - - - -   Regression: Fresh biomass, plant species 2   - - - - - - - - - -
The regression equation is plant2 = 81.9 + 46.7 Water

Predictor       Coef       Stdev    t-ratio        p
Constant       81.94       26.97       3.04    0.005
Water         46.739       6.030       7.75    0.000

s = 63.82       R-sq = 69.8%     R-sq(adj) = 68.6%
 - - - - - - - - - -   Regression: Fresh biomass, plant species 3   - - - - - - - - - -
The regression equation is plant3 = 33.0 + 25.4 Water

Predictor       Coef       Stdev    t-ratio        p
Constant       33.02       12.98       2.55    0.017
Water         25.423       2.901       8.76    0.000

s = 30.70       R-sq = 74.7%     R-sq(adj) = 73.7%
 - - - - - - - - - -   Regression: Fresh biomass, plant species 4   - - - - - - - - - -
The regression equation is plant4 = 15.7 + 23.6 Water

Predictor       Coef       Stdev    t-ratio        p
Constant       15.69       13.98       1.12    0.272
Water         23.641       3.127       7.56    0.000

s = 33.09       R-sq = 68.7%     R-sq(adj) = 67.5%
 - - - - - - - - - - -   Regression: Dry biomass, plant species 1   - - - - - - - - - - -
The regression equation is plant1 = 29.0 + 15.0 Water

Predictor       Coef       Stdev    t-ratio        p
Constant      28.971       6.033       4.80    0.000
Water         14.973       1.349      11.10    0.000

s = 14.28       R-sq = 82.6%     R-sq(adj) = 81.9%
 - - - - - - - - - - -   Regression: Dry biomass, plant species 2   - - - - - - - - - - -
The regression equation is plant2 = 31.7 + 13.4 Water

Predictor       Coef       Stdev    t-ratio        p
Constant      31.707       8.905       3.56    0.001
Water         13.365       1.991       6.71    0.000

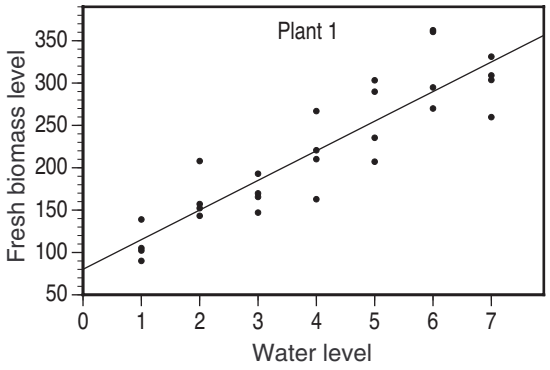
s = 21.07       R-sq = 63.4%     R-sq(adj) = 62.0%
 - - - - - - - - - - -   Regression: Dry biomass, plant species 3   - - - - - - - - - - -
The regression equation is plant3 = 18.4 + 8.44 Water

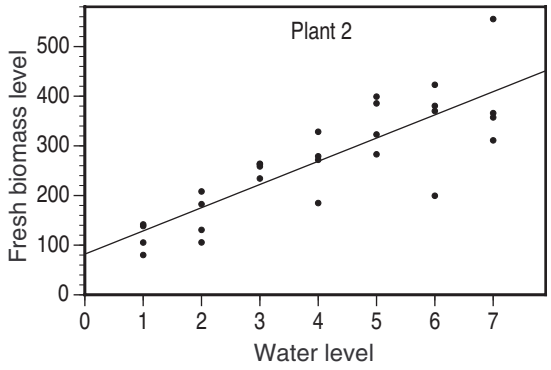
Predictor       Coef       Stdev    t-ratio        p
Constant      18.436       4.741       3.89    0.001
Water          8.442       1.060       7.96    0.000

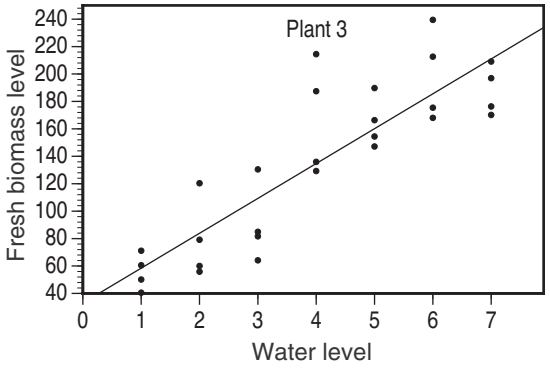
s = 11.22       R-sq = 70.9%     R-sq(adj) = 69.8%
 - - - - - - - - - - -   Regression: Dry biomass, plant species 4   - - - - - - - - - - -
The regression equation is plant4 = 10.3 + 7.20 Water

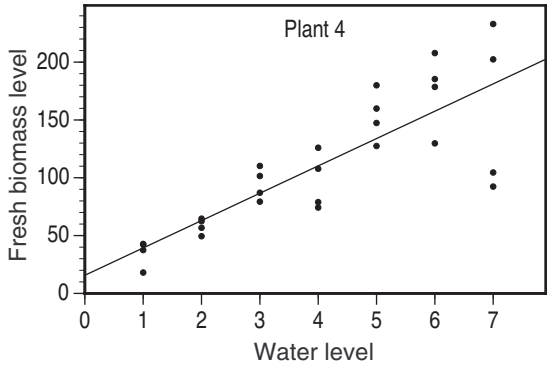
Predictor       Coef       Stdev    t-ratio        p
Constant      10.298       5.057       2.04    0.052
Water          7.197       1.131       6.36    0.000

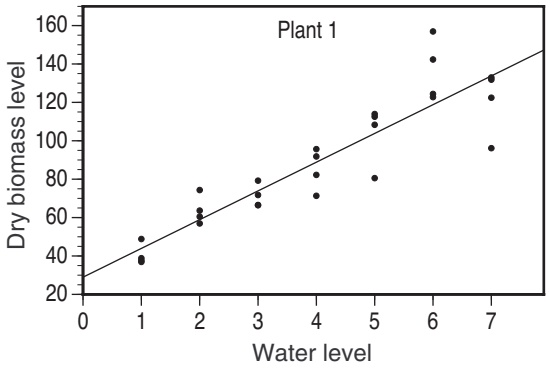
s = 11.97       R-sq = 60.9%     R-sq(adj) = 59.4%
```

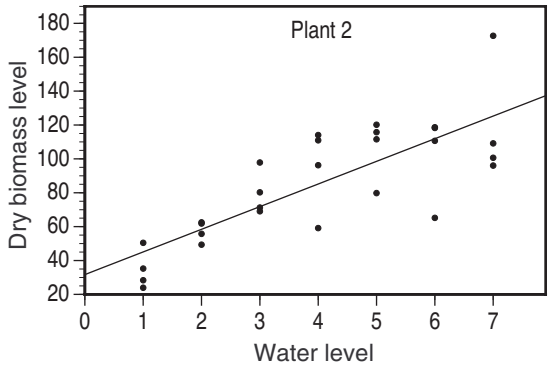

Plant 1
Fresh biomass level
Water level

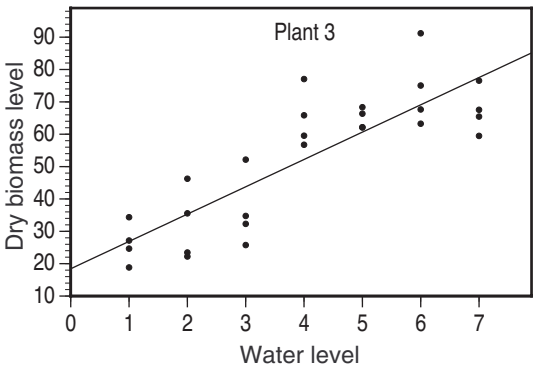

Plant 2
Fresh biomass level
Water level

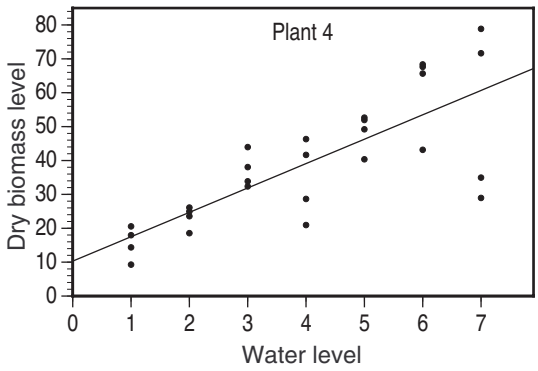

Plant 3
Fresh biomass level
Water level


Plant 4
Fresh biomass level
Water level


Plant 1
Dry biomass level
Water level

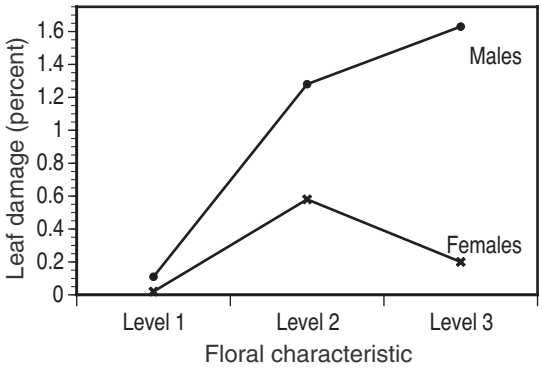

Plant 2
Dry biomass level
Water level


Plant 3
Dry biomass level
Water level


Plant 4
Dry biomass level
Water level

13.45. (a) With $I = 2$, $J = 3$, and $N = 180$, the numerator degrees of freedom are $I - 1$, $J - 1$, and $(I - 1)(J - 1)$, respectively, and the denominator degrees of freedom for all three tests is DFE $= N - IJ = 174$:

Source	df
Gender	1 and 174
Floral characteristic	2 and 174
Interaction	2 and 174



(b) Damage to males was higher for all characteristics. For males, damage was highest under characteristic level 3, while for females, the highest damage occurred at level 2. **(c)** Three of the standard deviations are at least half as large as the means. Because the response variable (leaf damage) had to be nonnegative, this suggests that these distributions are right-skewed; taking logarithms reduces the skewness.

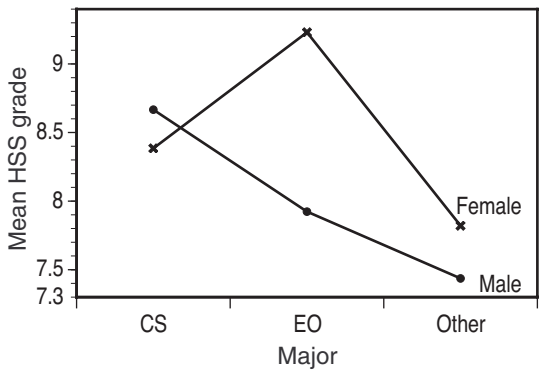
13.46. The table and plot of the means below suggest that, within a given gender, students who stay in the sciences have higher HSS grades than those who end up in the "Other" group. Males have a slightly higher mean in the CS group, but females have the edge in the other two. Normal quantile plots show no great deviations from Normality, apart from the granularity of the grades (most evident among women in EO). In the ANOVA, both main effects and the interaction are all significant. Residual analysis (not shown) shows that they are left-skewed.

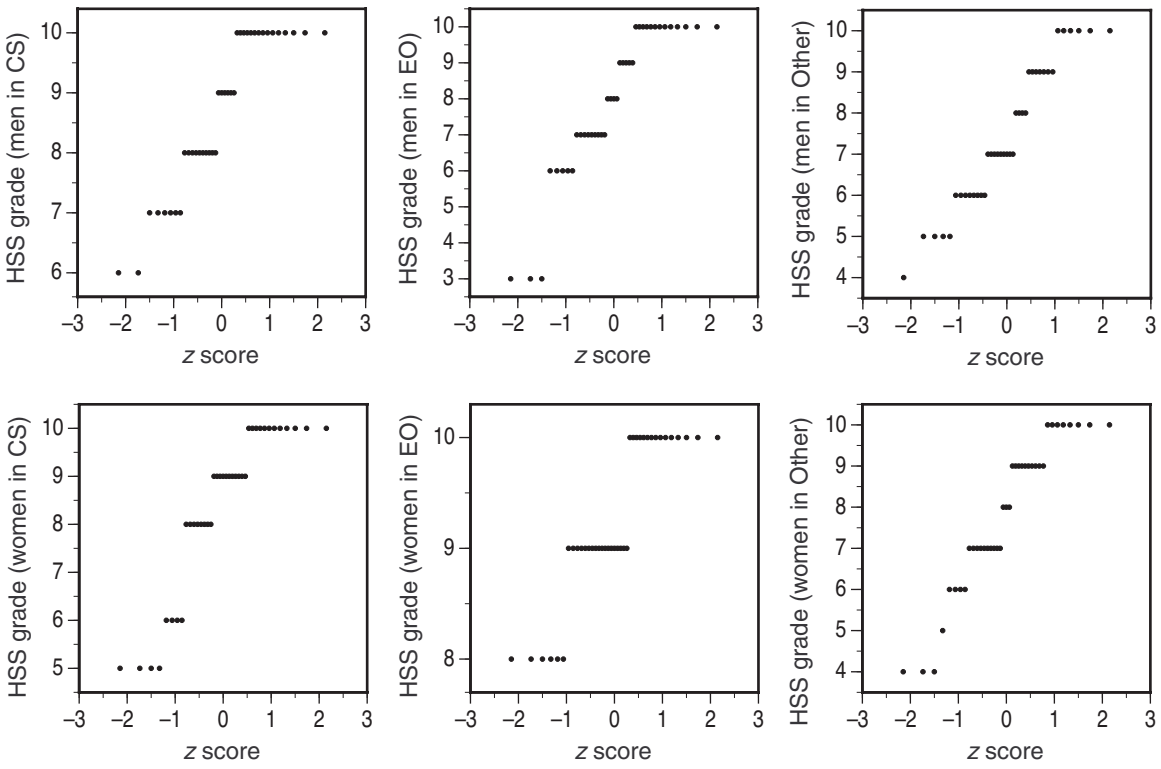
Minitab output

```
Analysis of Variance for HSS

Source       DF         SS         MS       F      P
Sex           1     12.927     12.927    5.06  0.025
Maj           2     44.410     22.205    8.69  0.000
Sex*Maj       2     24.855     12.427    4.86  0.009
Error       228    582.923      2.557
Total       233    665.115
```

		Major		
Gender		CS	EO	Other
Male	n	39	39	39
	$\bar{x}$	8.6667	7.9231	7.4359
	s	1.2842	2.0569	1.7136
Female	n	39	39	39
	$\bar{x}$	8.3846	9.2308	7.8205
	s	1.6641	0.7057	1.8046





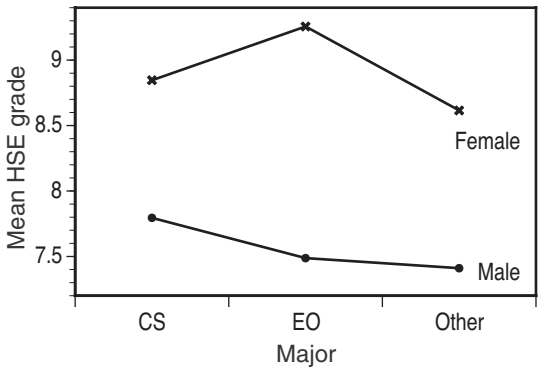
13.47. The table and plot of the means suggest that females have higher HSE grades than males. For a given gender, there is not too much difference among majors. Normal quantile plots show no great deviations from Normality, apart from the granularity of the grades (most evident among women in EO). In the ANOVA, only the effect of gender is significant. Residual analysis (not shown) reveals some causes for concern; for example, the variance does not appear to be constant.

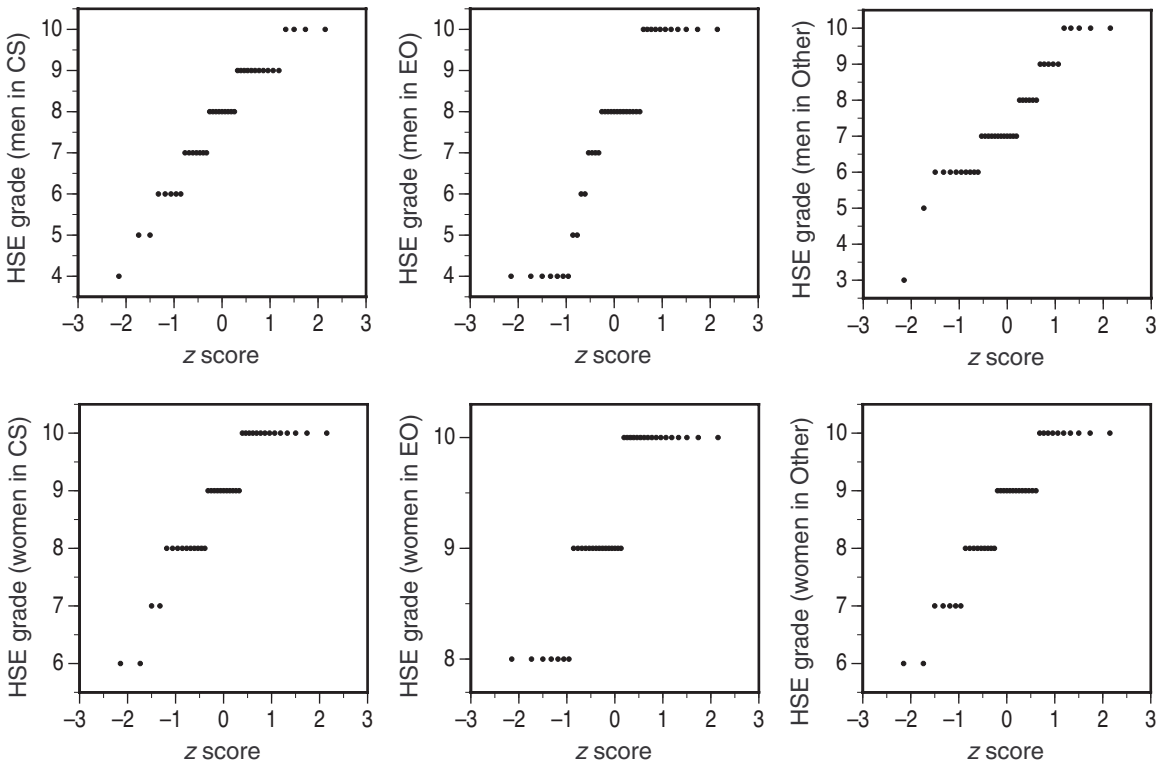
Minitab output

```
Analysis of Variance for HSE

Source       DF         SS         MS        F      P
Sex           1    105.338    105.338    50.32  0.000
Maj           2      5.880      2.940     1.40  0.248
Sex*Maj       2      5.573      2.786     1.33  0.266
Error       228    477.282      2.093
Total       233    594.073
```

		Major	
Gender	CS	EO	Other
Male	$n = 39$	39	39
	$\bar{x} = 7.7949$	7.4872	7.4103
	$s = 1.5075$	2.1505	1.5681
Female	$n = 39$	39	39
	$\bar{x} = 8.8462$	9.2564	8.6154
	$s = 1.1364$	0.7511	1.1611





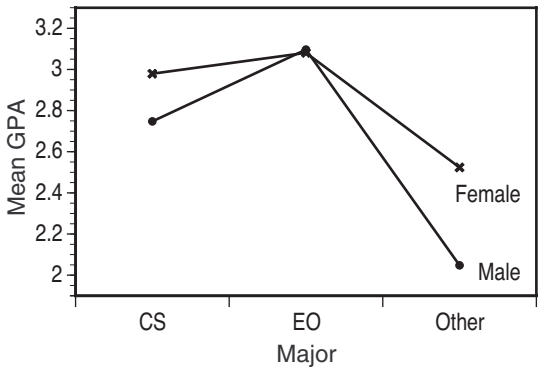
13.48. The table and plot of the means suggest that students who stay in the sciences have higher mean GPAs than those who end up in the "Other" group. Both genders have similar mean GPAs in the EO group, but in the other two groups, females perform better. Normal quantile plots show no great deviations from Normality, apart from a few low outliers in the two EO groups. In the ANOVA, sex and major are significant, while there is some (not quite significant) evidence for the interaction.

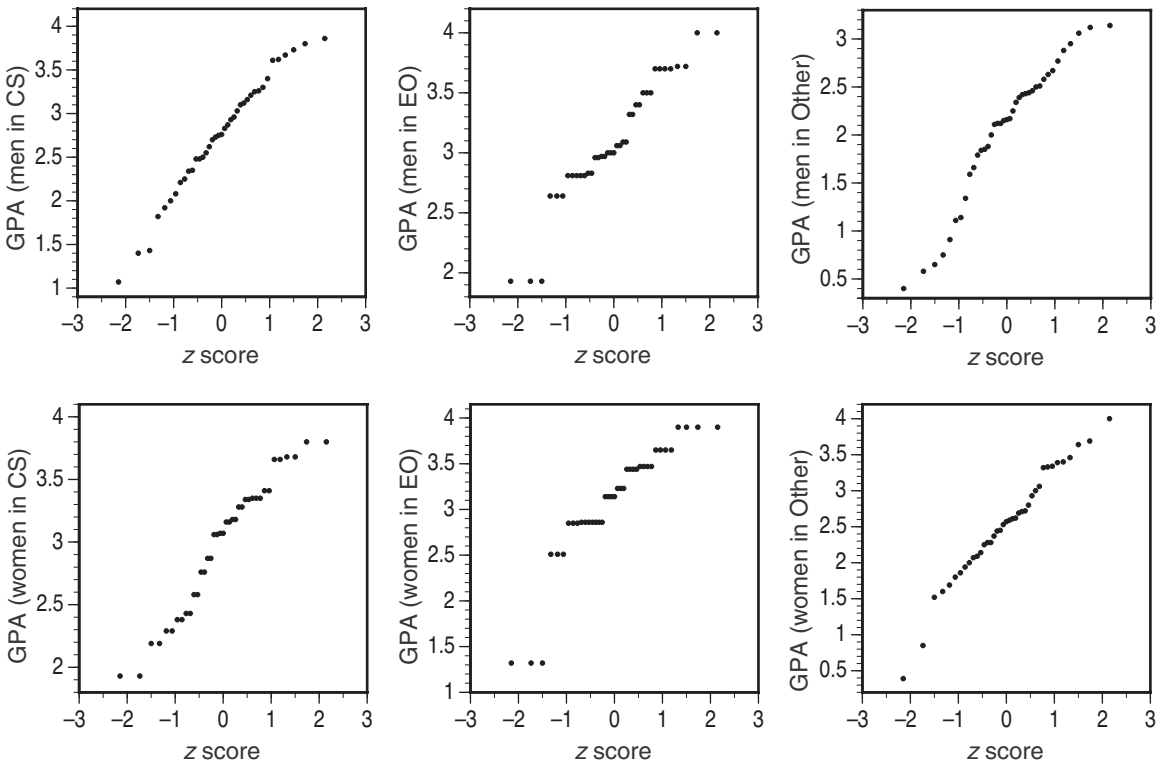
Minitab output

```
Analysis of Variance for GPA

Source       DF         SS         MS       F      P
Sex           1     3.1131     3.1131    7.31  0.007
Maj           2    26.7591    13.3795   31.42  0.000
Sex*Maj       2     2.3557     1.1779    2.77  0.065
Error       228    97.0986     0.4259
Total       233   129.3265
```

		Major		
Gender	CS	EO	Other	
Male	$n = 39$	39	39	
	$\bar{x} = 2.7474$	3.0964	2.0477	
	$s = 0.6840$	0.5130	0.7304	
Female	$n = 39$	39	39	
	$\bar{x} = 2.9792$	3.0808	2.5236	
	$s = 0.5335$	0.6481	0.7656	





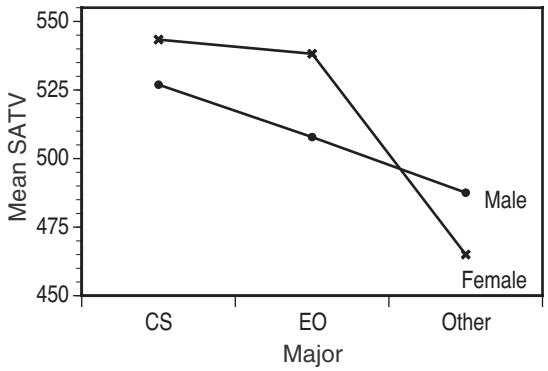
13.49. The table and plot of the means below suggest that students who stay in the sciences have higher mean SATV scores than those who end up in the "Other" group. Female CS and EO students have higher scores than males in those majors, but males have the higher mean in the Other group. Normal quantile plots suggest some right-skewness in the "Women in CS" group and also some non-Normality in the tails of the "Women in EO" group. Other groups look reasonably Normal. In the ANOVA table, only the effect of major is significant.

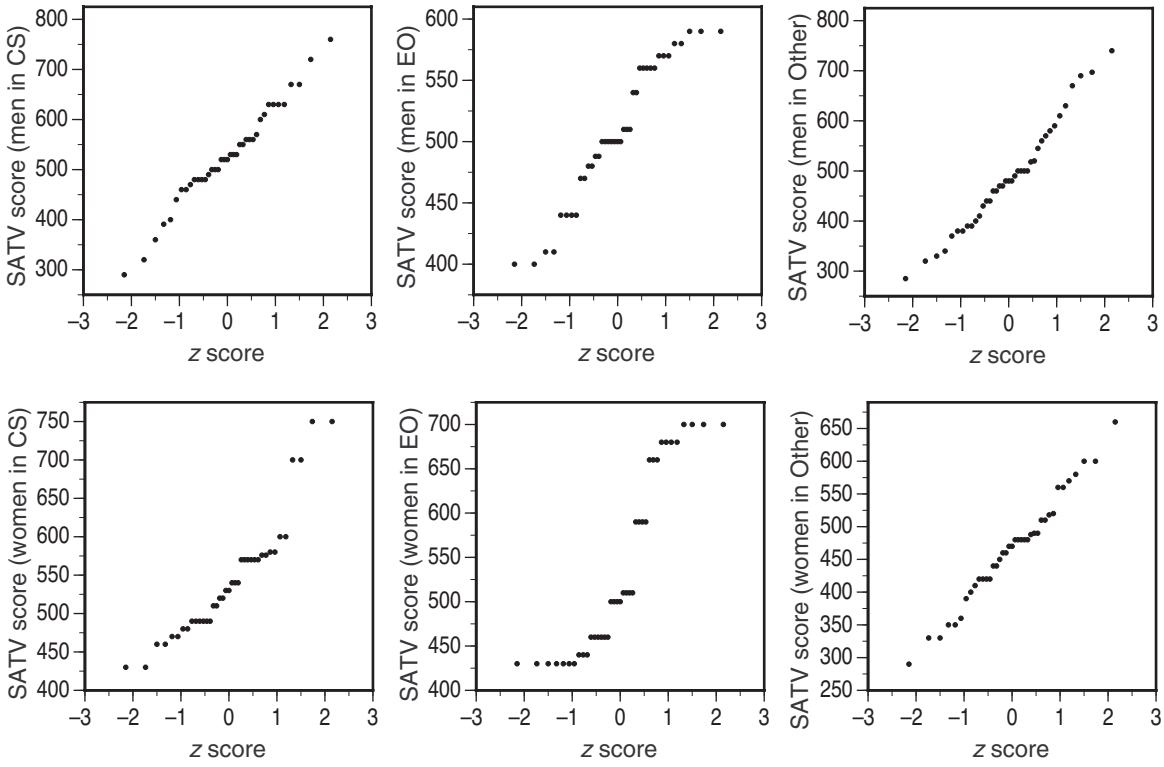
Minitab output

```
Analysis of Variance for SATV

Source       DF          SS          MS        F      P
Maj           2      150723       75362     9.32  0.000
Sex           1        3824        3824     0.47  0.492
Maj*Sex       2       29321       14661     1.81  0.166
Error       228     1843979        8088
Total       233     2027848
```

Gender	Major: CS	EO	Other
Male	$n =$ 39	39	39
	$\bar{x} =$ 526.949	507.846	487.564
	$s =$ 100.937	57.213	108.779
Female	$n =$ 39	39	39
	$\bar{x} =$ 543.385	538.205	465.026
	$s =$ 77.654	102.209	82.184




SATV score (men in CS)
800
700
600
500
400
300
–3 –2 –1 0 1 2 3
z score
SATV score (men in EO)
600
550
500
450
400
–3 –2 –1 0 1 2 3
z score
SATV score (men in Other)
800
700
600
500
400
300
–3 –2 –1 0 1 2 3
z score
SATV score (women in CS)
750
700
650
600
550
500
450
400
–3 –2 –1 0 1 2 3
z score
SATV score (women in EO)
700
650
600
550
500
450
400
–3 –2 –1 0 1 2 3
z score
SATV score (women in Other)
650
600
550
500
450
400
350
300
250
–3 –2 –1 0 1 2 3
z score