Chapter 9: More Complicated Experimental Designs

9.1 a. The F-test from the ANOVA table tests 2-sided alternatives:
Test $H_o : \mu_{Attend} = \mu_{DidNot}$ vs $H_o : \mu_{Attend} \neq \mu_{DidNot}$
The ANOVA table is given here:

Source	DF	SS	MS	F	p-value
Pair	5	1319.42	263.88		
Treatment	1	420.08	420.08	71.40	0.0001
Error	5	29.42	5.88		
Total	11	1768.92			

Reject H_o and conclude there is significant evidence that the mean scores of students attending Head Start are significantly different from the mean scores of students who do not attend Head Start.

b. RE(RCB,CR)$=\frac{(b-1)MSB+b(t-1)MSE}{(bt-1)MSE} = \frac{(6-1)(263.88)+(6)(2-1)(5.88)}{((6)(2)-1)(5.88)} = 20.94 \Rightarrow$
It would take approximately 21 times as many observations (126) per treatment in a completely randomized design to achieve the same level of precision in estimating the treatment means as was accomplished in the randomized complete block design.

9.2 The F-test from the ANOVA table tests 2-sided alternatives

$H_o : \mu_{Attend} = \mu_{DidNot}$ vs $H_o : \mu_{Attend} \neq \mu_{DidNot}$.

The paired t-test can test one-sided alternatives:

$H_o : \mu_{Attend} \leq \mu_{DidNot}$ vs $H_o : \mu_{Attend} \geq \mu_{DidNot}$

$\bar{y}_{Attend} = 78.33$ $\bar{y}_{DidNot} = 66.50$

$t = \frac{\bar{D}}{S_D/\sqrt{n}} = \frac{11.83}{3.43/\sqrt{6}} = 8.45$

p-value $= Pr(t_5 \geq 8.45) = 0.0002 \Rightarrow$

Yes, there is significant evidence that the mean scores attending Head Start are greater than the mean scores of students who do not attend.

Note $(8.45)^2 = 71.40$

9.3 a. Blocks are Investigators and Treatments are Mixtures

b. Randomly assign the four Mixtures to the each of the Investigators

c. The randomized complete block design guarantees that each investigator measures each of the four mixtures, whereas in a completely randomized design, it is possible that some of the investigators may not measure some of the mixtures. This may cause a bias towards some of the mixtures if a particular investigator tends to always give high readings no matter which mixture is measured.

9.4 a. $y_{ij} = \mu + \beta_j + \alpha_i + \epsilon_{ij}$; $i = 1, 2, 3, 4$ $j = 1, 2, 3, 4, 5$

y_{ij} is measurement of jth investigator on ith mixture

α_i is the ith mixture effect

β_j is the jth investigator effect

b. $\hat{\mu} = \bar{y}_{..} = 2463.75$

$\hat{\alpha}_1 = \bar{y}_{1.} - \bar{y}_{..} = 2351.0 - 2463.75 = -112.75$, $\hat{\alpha}_2 = \bar{y}_{2.} - \bar{y}_{..} = 2653.2 - 2463.75 = 189.45$

$\hat{\alpha}_3 = \bar{y}_{3.} - \bar{y}_{..} = 2444.2 - 2463.75 = -19.55$, $\hat{\alpha}_4 = \bar{y}_{4.} - \bar{y}_{..} = 2406.6 - 2463.75 = -57.15$

$\hat{\beta}_1 = \bar{y}_{.1} - \bar{y}_{..} = 2462.5 - 2463.75 = -1.25$, $\hat{\beta}_2 = \bar{y}_{.2} - \bar{y}_{..} = 2468.25 - 2463.75 = 4.5$

$\hat{\beta}_3 = \bar{y}_{.3} - \bar{y}_{..} = 2469.25 - 2463.75 = 5.5$, $\hat{\beta}_4 = \bar{y}_{.4} - \bar{y}_{..} = 2456.0 - 2463.75 = -7.75$

$\hat{\beta}_5 = \bar{y}_{.5} - \bar{y}_{..} = 2462.75 - 2463.75 = -1$

c. $F = 1264.73$ with p-value $< 0.0001 \Rightarrow$

Reject $H_o : \mu_1 = \mu_2 = \mu_3 = \mu_4$ and conclude there is significant evidence of a difference in the means for the four mixtures.

d. Mixture 2 with the highest mean response would appear to be possibly the best mixture.

A multiple comparison procedure could be used to confirm that the other three mixtures have significantly lower means.

e. RE(RCB,CR)$= \frac{(b-1)MSB + b(t-1)MSE}{(bt-1)MSE} = \frac{(5-1)(113.12)+(5)(4-1)(68.86)}{((5)(4)-1)(68.86)} = 1.14 \Rightarrow$

It would take 1.14 times as many observations (approximately 6) per treatment in a completely randomized design to achieve the same level of precision in estimating the treatment means as was accomplished in the randomized complete block design.

9.5 a. $y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}$; $i = 1, 2, 3$, $j = 1, 2, 3, 4, 5, 6, 7$

y_{ij} is score on test of jth subject hearing the ith music type

α_i is the ith music type effect

β_j is the jth subject effect

$\hat{\mu} = 21.33$, $\hat{\alpha}_1 = -0.47$, $\hat{\alpha}_2 = -1.19$, $\hat{\alpha}_3 = 1.67$

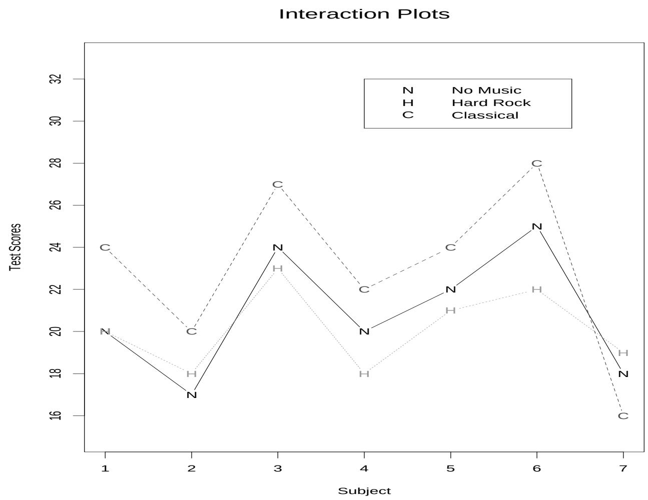
$\hat{\beta}_1 = 0$, $\hat{\beta}_2 = -3$, $\hat{\beta}_3 = 3.33$, $\hat{\beta}_4 = -1.33$, $\hat{\beta}_5 = 1$, $\hat{\beta}_6 = 3.67$, $\hat{\beta}_7 = -3.67$

b. $F = \frac{SS_{TRT}/df_{TRT}}{SS_{Error}/df_{Error}} = \frac{30.952/2}{28.38/12} = 6.54$ with df=2,12.

Therefore, p-value $= Pr(F_{2,12} \geq 6.54) = 0.0120 \Rightarrow$ Reject $H_o : \mu_1 = \mu_2 = \mu_3$.

We thus conclude that there is significant evidence of a difference in mean typing scores for the three types of music.

c. An interaction plot of the data is given here:



Based on the interaction plot, the additive model may be inappropriate because there is some crossing of the three lines. However, the plotted points are means of a single observation and hence may be quite variable in their estimation of the population means μ_{ij}. Thus, exact parallelism is not required in the data plots to ensure the validity of the additive model.

d. $t = 3, b = 7 \Rightarrow RE(RCB, CR) = \frac{(7-1)(24.889)+(7)(3-1)(2.365)}{((7)(3)-1)(2.365)} = 3.86 \Rightarrow$
It would take 3.86 times as many observations (approximately 27) per treatment in a completely randomized design to achieve the same level of precision in estimating the treatment means as was accomplished in the randomized complete block design. Since RE was much larger than 1, we would conclude that the blocking was effective.

9.6 The model conditions appear to be satisfied:

The normal probability plots and box plots of the residuals do not indicate nonnormality.

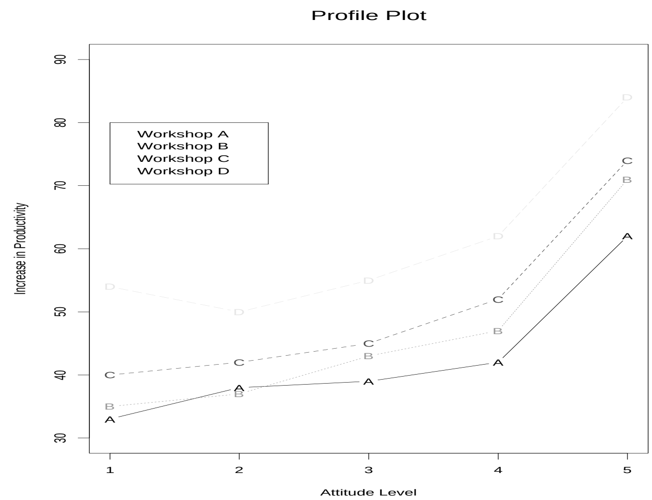
Plot of residuals versus estimated mean does not indicate nonconstant variance

Interaction plot indicates a potential interaction between subjects and type of music, but the indications are fairly weak.

9.7 a. $y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}; \quad i = 1, 2, 3, 4 \quad j = 1, 2, 3, 4, 5$
y_{ij} is the increase in productivity of worker having jth level of attitude and attending workshop type ith
α_i is the ith workshop type effect
β_j is the jth attitude effect
$\hat{\mu} = 50.25, \quad \hat{\alpha}_1 = -7.45, \quad \hat{\alpha}_2 = -3.65, \quad \hat{\alpha}_3 = 0.35, \quad \alpha_4 = 10.75$
$\hat{\beta}_1 = -9.75, \quad \hat{\beta}_2 = -8.5, \quad \hat{\beta}_3 = -4.75, \quad \hat{\beta}_4 = 0.5, \quad \hat{\beta}_5 = 22.5$

b. $F = \frac{SS_{TRT}/df_{TRT}}{SS_{Error}/df_{Error}} = \frac{922.55/3}{55.7/12} = 114.12$ with df=3,12.
Therefore, p-value $= Pr(F_{3,12} \geq 114.12) < 0.0001 \Rightarrow$ Reject $H_o : \mu_1 = \mu_2 = \mu_3 = \mu_4$.
We thus conclude that there is significant evidence of a difference in the mean increase in productivity for the four types of workshops.

c. A profile plot of the data is given here:



Based on the profile plot, the additive model appears to be appropriate because the four lines are relatively parallel. Note further that the plotted points are means of a single observation and hence may be quite variable in their estimation of the population means μ_{ij}. Thus, exact parallelism is not required in the data plots to ensure the validity of the additive model.

d. $t = 4, b = 5 \Rightarrow RE(RCB, CR) = \frac{(5-1)(696.375)+(5)(4-1)(4.6417)}{((5)(4)-1)(4.6417)} = 32.37 \Rightarrow$
It would take 32.37 times as many observations (approximately 162) per treatment in a completely randomized design to achieve the same level of precision in estimating the treatment means as was accomplished in the randomized complete block design. Since RE was much larger than 1, we would conclude that the blocking was effective.

9.8 The model conditions appear to be satisfied:

The normal probability plots and box plots of the residuals do not indicate nonnormality.

Plot of residuals versus estimated mean does not indicate nonconstant variance

Interaction plot indicates that the additive model (no interaction between attitude and type of workshop) is valid.

9.9 a. $y_{ijk} = \mu + \alpha_k + \beta_i + \gamma_j + \epsilon_{ijk}; \quad i, j, k = 1, 2, 3, 4;$
where y_{ijk} is the dry weight of a watermellon plant grown in Row i and Column j receiving Treatment k.

α_k is the effect of the kth Treatment on dry weight

β_i is the effect of the ith Row on dry weight

γ_j is the effect of the jth Column on dry weight

b. The Row, Column, and Treatment Means are given here:

Level	1	2	3	4
Row Mean $\bar{y}_{i..}$	1.53	1.5475	1.545	1.5475
Column Mean $\bar{y}_{.j.}$	1.5625	1.575	1.505	1.5275
Treatment Mean $\bar{y}_{.k.}$	1.7375	1.685	1.4225	1.325

The overall mean is $\bar{y}_{...} = 1.5425$
The parameter estimates are given here:
$\hat{\mu} = 1.5425$, $\hat{\beta}_1 = -.0125$, $\hat{\beta}_2 = .005$, $\hat{\beta}_3 = 0.0025$, $\beta_4 = .005$
$\hat{\gamma}_1 = .02$, $\hat{\gamma}_2 = .0325$ $\hat{\gamma}_3 = -.0375$, $\hat{\alpha}_4 = -.015$
$\hat{\alpha}_1 = .195$, $\hat{\alpha}_2 = .1425$ $\hat{\alpha}_3 = -.12$, $\hat{\alpha}_4 = -.2175$

$F = \frac{SS_{Trt}/df_{Trt}}{SS_{Error}/df_{Error}} = \frac{.48015/3}{.00075/6} = 1280.4$ with $df = 3, 6 \Rightarrow$ p-value < 0.0001
Reject $H_o : \mu_1 = \mu_2 - \mu_3 = \mu_4$ and conclude there is significant evidence that the four treatments have different mean dry weights.

9.10 a. $y_{ij} = \mu + \alpha_k + \beta_i + \gamma_j + \epsilon_{ij}$; $i, j, k = 1, 2, 3, 4$;
where y_{ij} is the mileage of a Driver i in Car Model j.
α_k is the effect of the kth Gasoline Blend on mileage
β_i is the effect of the ith Driver on mileage
γ_j is the effect of the jth Car Model on mileage

b. The Row, Column, and Treatment Means are given here:

Level	1	2	3	4
Driver Mean $\bar{y}_{i.}$	23.05	21.08	22.40	22.45
Model Mean $\bar{y}_{.j}$	14.08	31.55	17.50	25.85
Blend Mean $\bar{y}_k$	22.50	24.98	18.05	23.45

The overall mean is $\bar{y}_{...} = 22.245$
The parameter estimates are given here:
$\hat{\mu} = 22.245$, $\hat{\alpha}_1 = .255$, $\hat{\alpha}_2 = 2.735$ $\hat{\alpha}_3 = -4.195$, $\hat{\alpha}_4 = 1.205$
$\hat{\beta}_1 = .805$, $\hat{\beta}_2 = -1.165$, $\hat{\beta}_3 = .155$, $\beta_4 = .205$
$\hat{\gamma}_1 = -8.165$, $\hat{\gamma}_2 = 9.305$ $\hat{\gamma}_3 = -4.745$, $\hat{\alpha}_4 = 3.605$

c. The ANOVA table is given here:

Source	DF	SS	MS	F	p-value
Blend	3	106.272	35.424	8.22	0.015
Driver	3	8.332	2.777		
Car Model	3	755.372	251.791		
Error	6	25.864	4.311		
Total	15	895.839			

p-value $= 0.015 < 0.05 \Rightarrow$ Reject H_o and conclude there is significant evidence that there is a difference in the mean mileage of the four blends of gasoline.

d. The blend with the highest mileage is Blend B. However, there is very little difference between the sample means of Blends A, B, and D. In fact, the estimate $SE(\hat{\mu}_k) = \hat{\sigma}_\epsilon/\sqrt{n} = 4.311/\sqrt{4} = 2.16$. Therefore, Blends A, B, and D differ by only about one SE and hence would not be significantly different.

e. RE(LS,CR)=$\frac{MSR+MSC+(t-1)MSE}{(t+1)MSE} = \frac{2.777+251.791+(4-1)(4.311)}{((4+1)(4.311)} = 12.41 \Rightarrow$

It would take 12.41 times as many observations (approximately 50) per treatment in a completely randomized design to achieve the same level of precision in estimating the treatment means as was accomplished in the Latin square design.

f. In future studies, this type of study could be run as a RCB design using Car Model as the blocking variable because the four drivers had very little variation in their mileage.

9.11 a. Similar results.

b. The Boxplot and normal probability plot do not indicate a deviation from a normal distribution for the residuals.

The plot of Residuals vs Pred do not indicate a deviation from the constant variance condition.

Based on these plots, there does not indicate any deviations from the model conditions.

9.12 a. Completely randomized design with a 3x2 factorial treatment structure and 10 reps.

b. $y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijk}; \quad i = 1, 2, 3; \quad j = 1, 2; \quad k = 1, \cdots, 10;$

where y_{ijk} is the attention span of the kth child of Age i viewing Product j.

α_i is the effect of the ith Age on attention span

β_j is the effect of the jth Product on attention span

$\alpha\beta_{ij}$ is the interaction effect of the ith Age and jth Product on attention span

c. The treatment means are given here:

Treatment	A_1P_1	A_2P_1	A_3P_1	A_1P_2	A_2P_2	A_3P_2
Sample Mean	22.9	19.6	21.9	23.1	30.5	45.6
Factor Level	A_1	A_2	A_3		P_1	P_2
Sample Mean	23.0	25.05	33.75		21.47	33.07

$\hat{\mu} = 27.27, \quad \hat{\alpha}_1 = -4.27, \quad \hat{\alpha}_2 = -2.22, \quad \hat{\alpha}_3 = 6.48$

$\hat{\beta}_1 = -5.8, \quad \hat{\beta}_2 = 5.8$

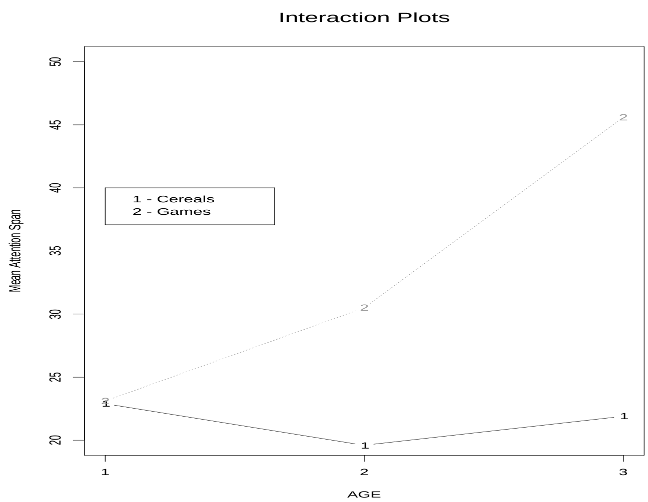
$\hat{\alpha}\beta_{11} = 5.7, \quad \hat{\alpha}\beta_{21} = .35, \quad \hat{\alpha}\beta_{31} = -6.05$

$\hat{\alpha}\beta_{12} = -5.7, \quad \hat{\alpha}\beta_{22} = -.35, \quad \hat{\alpha}\beta_{32} = 6.05$

d. The ANOVA table is given here:

Source	DF	SS	MS	F	p-value
Age	2	1303.0	651.5	4.43	0.017
Product	1	2018.4	2018.4	13.72	0.001
Interaction	2	1384.3	692.1	4.70	0.013
Error	54	7944	147.1		
Total	59	12649.7			

9.13 a. A profile plot of the data is given here:



The profile plot indicates an increasing effect of Product Type as Age increases.

b. The p-value for the interaction term is 0.013. There is significant evidence of an interaction between the factors Age and Product Type. Thus, the size of the difference between mean attention span of children viewing breakfast cereals and viewing video games would be different for the three age groups. From the profile plots, the estimated mean attention span for video games is larger than for breakfast cereals, with the size of the difference becoming larger as age increases.

9.14 a. Similar results are obtained.

b. The residuals in the normal probability plot appear to fall very close to a straight line and hence we can conclude there is not evidence that the residuals have a non-normal distribution.

The plot of the residuals versus Fitted Value appear to have a consistent width across the fitted values. The condition of nonconstant variance does not appear to be violated.

9.15 There are 15 treatments consisting of the 3 levels of Factor A combined with the 5 levels of Factor B. These 15 treatments will be randomly assigned to 15 experimental units in each of the three blocks as seen in the following diagram:

	Block 1			Block 2			Block 3		
	Factor B			Factor B			Factor B		
Factor A	B1	B2	B3	B1	B2	B3	B1	B2	B3
A1	x	x	x	x	x	x	x	x	x
A2	x	x	x	x	x	x	x	x	x
A3	x	x	x	x	x	x	x	x	x
A4	x	x	x	x	x	x	x	x	x
A5	x	x	x	x	x	x	x	x	x

Source	DF	SS	MS	F	p-value
Treatment	14	SST	MST		
Factor A	2	SSA	MSA	MSA/MSE	
Factor B	4	SSB	MSB	MSB/MSE	
Interaction	8	SSAB	MSAB	MSAB/MSE	
Blocks	2	SSBL	MSBL		
Error	28	SSE	MSE		
Total	44	SSTOT			

9.16 There are 24 treatments consisting of the 2 levels of Factor A combined with the 4 levels of Factor B with 3 levels of Factor C. These 24 treatments will be randomly assigned to 24 experimental units in each of the two blocks as seen in the following diagram:

	Block 1						Block 2					
	A1			A2			A1			A2		
Factor B	C1	C2	C3	C1	C2	C3	C1	C2	C3	C1	C2	C3
B1	x	x	x	x	x	x	x	x	x	x	x	x
B2	x	x	x	x	x	x	x	x	x	x	x	x
B3	x	x	x	x	x	x	x	x	x	x	x	x
B4	x	x	x	x	x	x	x	x	x	x	x	x

Source	DF	SS	MS	F	p-value
Treatments	23	SST	MST		
Factor A	1	SSA	MSA	MSA/MSE	
Factor B	3	SSB	MSB	MSB/MSE	
A*B	3	SSAB	MSAB	MSAB/MSE	
Factor C	2	SSB	MSC	MSC/MSE	
A*C	2	SSAB	MSAC	MSAC/MSE	
B*C	6	SSAB	MSBC	MSBC/MSE	
A*B*C	6	SSAB	MSABC	MSABC/MSE	
Blocks	1	SSBL	MSBL		
Error	23	SSE	MSE		
Total	47	SSTOT			

9.17 a. This is a randomized complete block design with the eight regions serving as the blocking variable and type of job serving as the treatment.

$y_{ij} = \mu + \alpha_i + \beta_j + \epsilon_{ij}; \quad i = 1, 2, 3, \quad j = 1, 2, 3, 4, 5, 6, 7, 8$
y_{ij} is starting salary in region j of job type i
α_i is the ith job type effect on starting salary
β_j is the jth region effect on starting salary

b. The Minitab output is given here:

```
General Linear Model: y versus Region, Group

Factor     Type Levels Values
Region    fixed      8 1 2 3 4 5 6 7 8
Group     fixed      3 P F I


Analysis of Variance for y, using Adjusted SS for Tests

Source     DF     Adj SS     Adj MS       F      P
Region      7     42.620      6.089   14.42  0.000
Group       2     79.491     39.745   94.16  0.000
Error      14      5.909      0.422
Total      23    128.020
```

Because the p-value $< 0.0001 < 0.05 = \alpha$, we can conclude there is significant evidence that the mean starting salary for the three groups of employees is different.

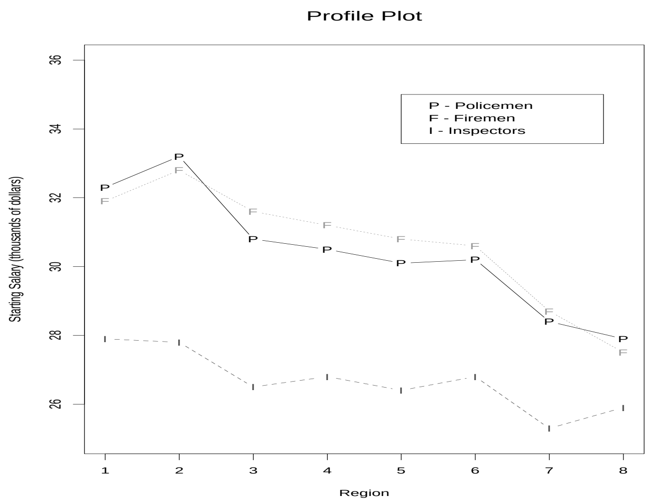
c. p-value < 0.0001

d. Using Tukey's W-procedure with $\alpha = 0.05, s_\epsilon^2 = MSE = .422, q_\alpha(t, df_{error}) = q_{.05}(3, 14) = 3.70 \Rightarrow$
$W = (3.70)\sqrt{\frac{.422}{8}} = .85 \Rightarrow$

Employee Group	Inspectors	Policemen	Firemen
Sample Mean	26.68	30.42	30.64
Tukey Grouping	a	b	b

Thus, Inspectors have significantly lower starting salaries than both Policemen and Firemen.

9.18 a. A profile plot of the data is given here:



Based on the profile plot, the additive model appears to be appropriate because the three lines are relatively parallel. Note further that the plotted points are means of a single observation and hence may be quite variable in their estimation of the population means μ_{ij}. Thus, exact parallelism is not required in the profile plots to ensure the validity of the additive model.

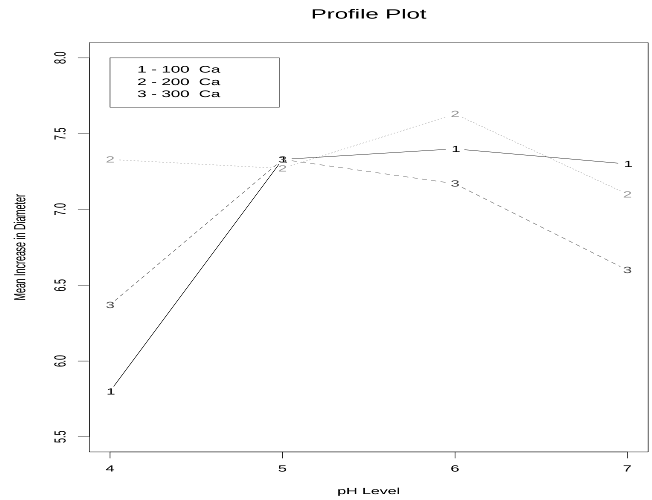
It would not be possible to test for an interaction between Region and Job Type because there is only one observation per Region-Job Type combination.

b. RE(RCB,CR)=$\frac{(b-1)MSB+b(t-1)MSE}{(bt-1)MSE} = \frac{(8-1)(6.089)+(8)(3-1)(.422)}{((8)(3)-1)(.422)} = 5.09 \Rightarrow$

It would take 5.09 times as many observations (approximately 41) per treatment in a completely randomized design to achieve the same level of precision in estimating the treatment means as was accomplished in the randomized complete block design.

c. Other possible important factors may be average salaries of all government employees in the region, education requirements for the position, etc.

9.19 a. The profile plot is given here:



There appears to be an interaction between Ca Rate and pH with respect to the increase in trunk diameters. At low pH value, a 300 level of Ca yields largest increase; whereas, at high pH value, a 100 level of Ca yields largest increase in trunk diameter.

b. A model for this experiment is given here:

$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijk}; \quad i = 1,2,3,4; \quad j = 1,2,3; \quad k = 1,2,3;$

where y_{ijk} is the increase in trunk diameter of the kth tree in soil having the ith pH level using the jth Ca Rate:

α_i is the effect of the ith pH level on diameter increase

β_j is the effect of the jth Ca Rate on diameter increase

$\alpha\beta_{ij}$ is the interaction effect of the ith pH level and jth Ca Rate on diameter increase.

c. This is a completely randomized 4x3 factorial experiment with Factor A: pH level, Factor B: Ca rate. There are 3 complete replications of the experiment. The AOV table is given here:

Source	DF	SS	MS	F	p-value
pH	3	4.461	1.487	21.94	0.0001
Ca	2	1.467	0.734	10.82	0.0004
Interaction	6	3.255	0.543	8.00	0.0001
Error	24	1.627	0.0678		
Total	35	10.810			

9.20 a. The p-value for pH by Ca interaction is p-value $< 0.0001 \Rightarrow$ there is signification evidence of an interaction between level of pH and rate of Ca on the mean increase in trunk diameter. Since the interaction is significant, the main effects do not have direct interpretation and hence the tests are not very meaningful.

b. Because there is a significant interaction between pH and Ca, any conclusions about the effect of Ca Rate on mean increase in trunk diameter will vary depending on the pH of the soil. When the pH=4, a 300 level of Ca appears to provide the greatest increase in trunk diameter. When pH=5 or 6, the mean trunk diameters are not very different for the three levels of Ca. When pH=7, Ca=100 provides the greatest increase in trunk diameter. In fact, we should conduct a Tukey's comparison on the three Ca levels at each level of pH: (See Exercise 15.26)

9.21 a. Using Tukey's W-procedure with $\alpha = 0.05, s_\epsilon^2 = MSE = .0678, q_\alpha(t, df_{error}) = q_{.05}(3, 24) = 3.53 \Rightarrow$

$W = (3.53)\sqrt{\frac{.0678}{3}} = 0.53 \Rightarrow$

		Ca Rate		
pH		100	200	300
4	Mean	5.80	7.33	6.37
	Grouping	a	c	b
5	Mean	7.33	7.27	7.33
	Grouping	a	a	a
6	Mean	7.40	7.63	7.17
	Grouping	a	a	a
7	Mean	7.30	7.10	6.60
	Grouping	b	ab	a

b. From the above table we observe that at pH=5, 6 there is not significant evidence of a difference in mean increase in diameter between the three levels of Ca. However, at pH=4, 7 there is significant evidence of a difference with Ca=300 yielding the largest increase at pH=4 and Ca=100 or 200 yielding the largest increase at pH=7. This illustrates the interaction between Ca and pH, i.e., the size of differences in the means across the levels of Ca depend on the level of pH.

9.22 a. The normality condition does not appear to be violated: Box plot is symmetric with no outliers; points in normal probability plot fall fairly close to a straight line.

The plot of residuals versus Estimated Treatment Means displays a somewhat decreasing variance as the Estimated Treatment Means increase.

b. The conditions do not appear to be violated hence no modifications in the data are required. If the pattern in the plot of the residuals versus Estimated Treatment Means was more distinct, a square root or log transformation may be required.

9.23 a. The design is a completely randomized 5x5 factorial experiment with 2 replications; Factor A is Exterior Temperature and Factor B is Pane Design.

A model for this experiment is given here:

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijk}; \quad i = 1, 2, 3, 4, 5; \quad j = 1, 2, 3, 4, 5; \quad k = 1, 2;$$

where y_{ijk} is the heat loss of the kth pane having the ith temperature level and the jth pane design:

α_i is the effect of the ith temperature on heat loss

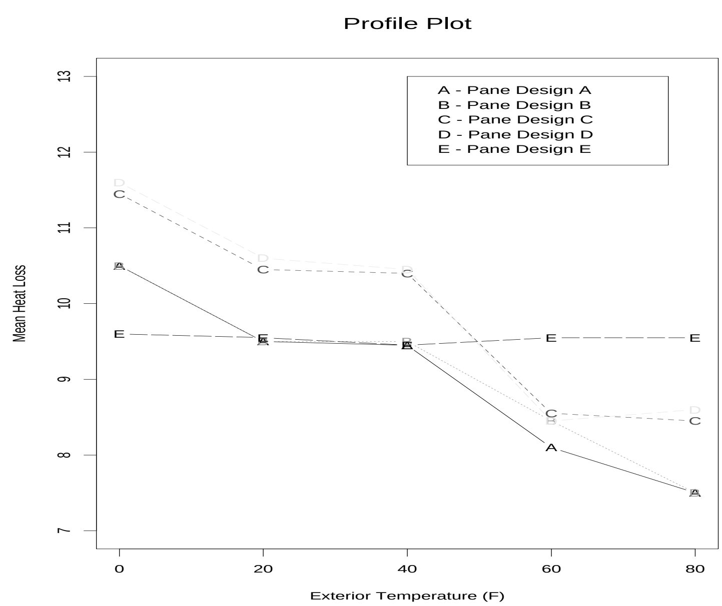
β_j is the effect of the jth pane desing on heat loss

$\alpha\beta_{ij}$ is the interaction effect of the ith temperature and jth pane design on heat loss

b. The test for an interaction between exterior temperature and pane design yields p-value = 0.0073 which would indicate significant evidence that an interaction exists. Therefore, the differences in mean heat loss between the five pane designs varies depending on the exterior temperature. The test of the main effect of pane design is not informative due to the significant interaction.

c. No, because of the significant interaction between pane design and exterior temperature. A profile plot is given here along with a table of the treatment means:

	Pane Design				
Temperature	A	B	C	D	E
0	10.50	10.50	11.45	11.60	9.60
20	9.50	9.50	10.45	10.60	9.55
40	9.45	9.50	10.40	10.45	9.45
60	8.10	8.45	8.55	8.45	9.55
80	7.50	7.50	8.45	8.60	9.55



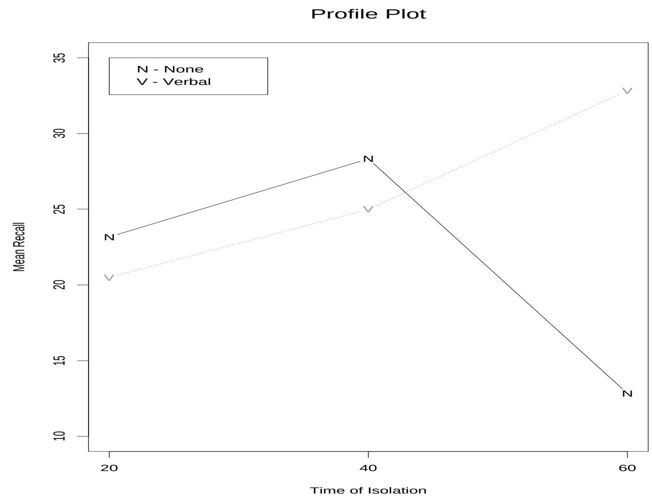
d. Using Tukey's W-procedure with $\alpha = 0.05, s_\epsilon^2 = MSE = .2312, q_\alpha(t, df_{error}) = q_{.05}(5, 25) = 4.16 \Rightarrow$
$W = (4.16)\sqrt{\frac{.2312}{2}} = 1.41 \Rightarrow$

		Pane Design				
		A	B	C	D	E
Temp=0	Mean	10.50	10.50	11.45	11.60	9.60
	Grouping	ab	ab	b	b	a
Temp=20	Mean	9.50	9.50	10.45	10.60	9.55
	Grouping	a	a	a	a	a
Temp=40	Mean	9.45	9.50	10.40	10.45	9.45
	Grouping	a	a	a	a	a
Temp=60	Mean	8.10	8.45	8.55	8.45	9.55
	Grouping	a	a	a	a	b
Temp=80	Mean	7.50	7.50	8.45	8.60	9.55
	Grouping	a	a	a	a	b

From the above table we observe that at the exterior temperatures of $20°F$ and $40°F$ there is not significant evidence of a difference in mean heat loss between the five pane designs. However, at the exterior temperatures of $60°F$ and $80°F$ pane design E has a significantly higher mean heat loss than the other four designs. At exterior temperature of $0°F$, there are two groups of pane designs relative to their mean heat loss. This illustrates the interaction between exterior temperature and pane design, i.e., the size of differences in mean heat loss between the five pane designs depends on the exterior temperature.

9.24 A profile plot is given here along with a table of the treatment means:

	Isolation Time		
Reinforcement	20	40	60
None	23.17	28.33	12.83
Verbal	20.50	25.00	32.83



The p-value for the test of an interaction between Level of Reinforcement and Time of Isolation is given on the output as p-value < 0.0001 which implies there is significant evidence of an interaction between the two factors. Therefore, in order to compare whether time of isolation has an effect on mean recall, it is necessary to compare the three levels of isolation at each level of reinforcement. Using Tukey's W-procedure with $\alpha = 0.05, s_\epsilon^2 = MSE = 15.78, q_\alpha(t, df_{error}) = q_{.05}(3, 30) = 3.49 \Rightarrow$

$W = (3.49)\sqrt{\frac{15.78}{6}} = 5.66 \Rightarrow$

Level of		Time of Isolation		
Reinforcement		20	40	60
None	Mean	23.17	28.33	12.83
	Grouping	b	b	a
Verbal	Mean	20.50	25.00	32.83
	Grouping	a	a	b

From the above table we observe that at a Level of Reinforcement of None the mean recall is significantly lower at an Isolation Time of 60 minutes but there is not evidence of a difference in mean recall between Isolation Times of 20 and 40. However, for Verbal Reinforcement the mean recall is significantly higher at an Isolation Time of 60 minutes but there is not evidence of a difference in mean recall between Isolation Times of 20 and 40. Thus, whether or not verbal reinforcement increases the mean recall depends on the time of isolation.

9.25 a. The experiment is run as three reps of a completely randomized design with a 2x4 factorial treatment structure. A model for the experiment is given here:

$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijk}; \quad i = 1, 2, 3, 4; \quad j = 1, 2; \quad k = 1, 2, 3;$
where y_{ijk} is the amount of active ingredient (or pH) of the kth vial having the ith storage time in laboratory jth:

α_i is the effect of the ith storage time on amount of active ingredient (or pH)

β_j is the effect of the jth laboratory on amount of active ingredient (or pH)

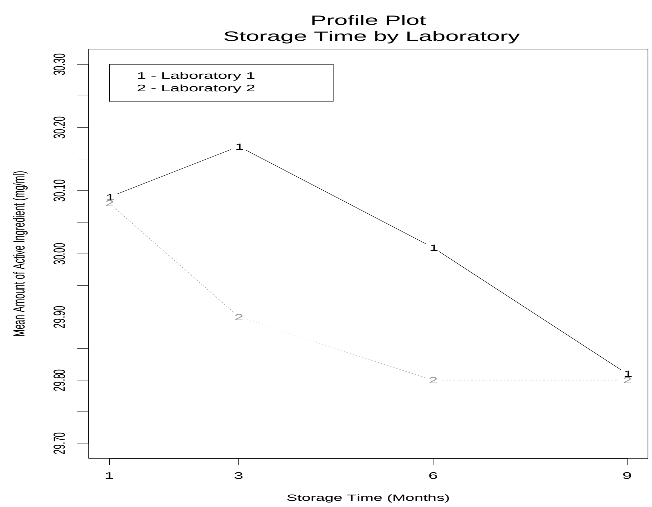
$\alpha\beta_{ij}$ is the interaction effect of the ith storage time and jth laboratory on amount of active ingredient (or pH)

b. The complete AOV table is given here:

Source	DF	SS	MS	F	p-value
Storage Time	3	SSA	SSA/3	MSA/MSE	
Laboratory	1	SSB	SSB/1	MSB/MSE	
Interaction	3	SSAB	SSAB/3	MSAB/MSE	
Error	16	SSE	SSE/16		
Total	23	SST			

9.26 The analysis for Amount of Active Ingredient:

There is significant evidence (p-value=0.0003) of an interaction between laboratory and time in storage. Therefore, a Tukey analysis of storage time will be conducted separately for each laboratory. The profile plot and Tukey analysis is given here:



Using Tukey's W-procedure with $\alpha = 0.05, s_\epsilon^2 = MSE = 0.0024458, q_\alpha(t, df_{error}) = q_{.05}(4, 16) = 4.05 \Rightarrow$

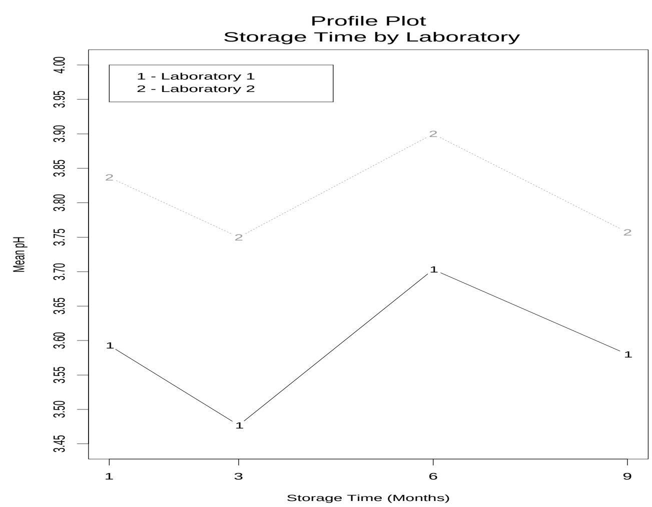
$W = (4.05)\sqrt{\frac{.0024458}{3}} = 0.116 \Rightarrow$

		Storage Time			
Laboratory		1	3	6	9
1	Mean	30.09	30.17	30.01	29.81
	Grouping	bc	c	b	a
2	Mean	30.08	29.90	29.80	29.80
	Grouping	b	a	a	a

From the above table we observe that for Laboratory 1, the mean amount of the active ingredient was lowest after a storage time of 9 months. There was not a significant difference in the mean amounts of the active ingredient for 1 and 3 months of storage and for 1 and 6 months of storage. For Laboratory 2, the results were somewhat different. There was a significant decline in the mean amount of the active ingredient after the first month of storage but the mean did not significantly change during the 3, 6, and 9 months of storage.

The analysis for pH:

There is not significant evidence (p-value=0.4038) of an interaction between laboratory and time in storage. Therefore, a Tukey analysis will be conducted for storage time and for laboratory. The profile plot and Tukey analysis is given here:



Using Tukey's W-procedure with $\alpha = 0.05, s_\epsilon^2 = MSE = .0027958, q_\alpha(t, df_{error}) = q_{.05}(3, 16) = 3.65 \Rightarrow$

$W = (3.65)\sqrt{\frac{.0027958}{6}} = .0788 \Rightarrow$

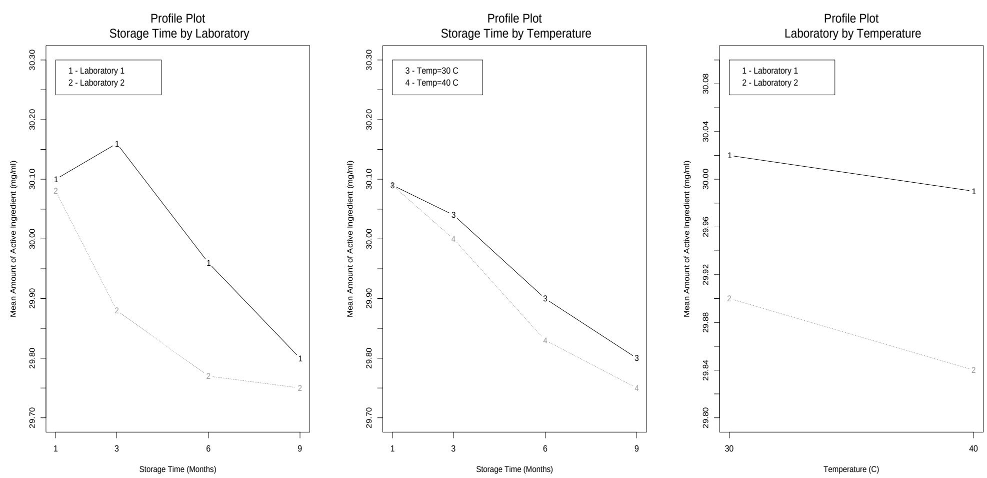
	Storage Time			
	1	3	6	9
Mean	3.715	3.614	3.802	3.669
Grouping	b	a	c	ab

The mean pH is lowest at storage times of 3 and 9 months and highest at 6 months.

The p-value for testing a difference in the mean pH between the two laboratories was less than 0.0001. Thus, there significant evidence that the vials have lower mean pH values at Laboratory 1 than at Laboratory 2 (3.588 vs 3.811).

9.27 a. With respect to Amount of Active Ingredient:

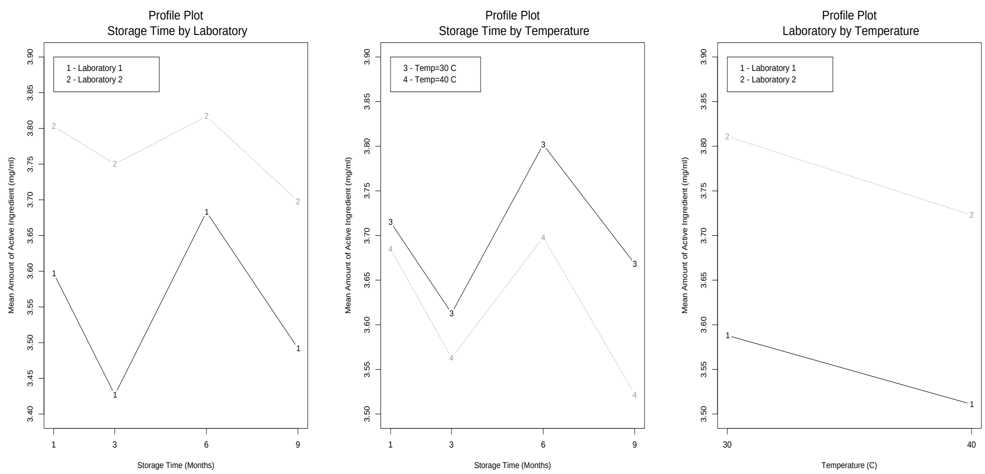
The profile plots for the two-way interactions are given here:



The Time*Lab*Temp interaction (p-value=.6914), Lab*Temp interaction (p-value=.3519), and Time*Temp interaction (p-value=.2817) were not significant. However, there was significant evidence (p-value=.0192) of a difference in the means for Amount of Active Ingredient at the two temperatures. There is a strong interaction between Time and Lab (p-value<.0001). Thus, comparisons of the means at the four storage times should be done separately for each Lab. The main effects of Lab and Time are not informative since these two factors have a significant interaction.

b. With respect to pH:

The profile plots for the two-way interactions are given here:



The Time*Lab*Temp interaction (p-value=.0621), Lab*Temp interaction (p-value = .7617), and Time*Temp interaction (p-value=.0686) were not significant. However, there was significant evidence (p-value $< .0001$) of a difference in the means for pH at the two temperatures. There is a strong interaction between Time and Lab (p-value=.0028). Thus, comparisons of the means at the four storage times should be done separately for each Lab. The main effects of Lab and Time are not informative since these two factors have a significant interaction.

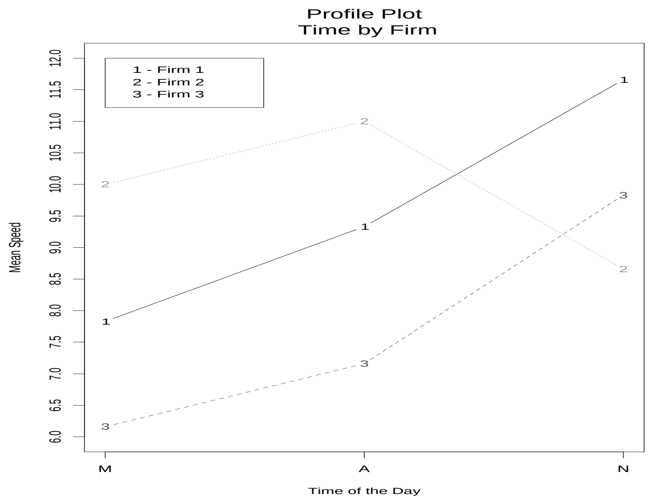
c. Because the interaction between storage time and laboratory is significant, the effects of storage time differ between the two labs. Lab 1 has highest mean pH after 6 months, followed by months 1,9 and 3 months, respectively. In contrast, Lab 2 shows the highest mean pH for 6 months and 1 month, followed by 3 and 9 months, respectively.

   Because the interaction between temperature and laboratory is not significant, the effects of temperature are the same for the two labs, with highest pH at 30°C.

d. Because the interaction between storage time and laboratory is significant, the effects of storage time differ between the two labs. Lab 1 has the highest mean active ingredient concentration after 3 months of storage time, followed by 1, 6, and 9 months, respectively. In contrast, Lab 2 shows its highest mean active ingredient concentration after 1 month of storage time, followed by 3, 6, and 9 months, respectively.

   Because the interaction between temperature and laboratory is not significant, the effects of storage time are the same for the two labs, with highest pH at 30°C..

9.28 a. The test for an interaction between Time and Firm would provide the evidence. The p-value for testing for an interaction is given as p-value=0.0016 which indicates that there is significant evidence of an interaction between Time and Firm. A profile plot will display the nature of the interaction:



b. No firm is uniformly the best; in the morning Firm 3 is the best; in the afternoon Firm 2 is the best; and at night Firm 1 is the best.

If the average performance over the three time periods is considered, then confidence intervals could be constructed for each firms mean elapsed time. If the confidence intervals overlap, then there is not a significant difference in the mean elapsed time of the firms. In order for the overall level of confidence to be maintained for the three intervals, it is necessary to divide α by 3 prior to dividing by 2:

Simultaneous 95% C.I.'s for the mean elapsed times μ_i's are constructed here:

$t_{.05/6,45} = 2.487$ $\hat{\sigma} = \sqrt{MSE} = 1.7847$, C.I. is $\bar{y}_{i.} \pm t_{.0083,45}\frac{\hat{\sigma}}{\sqrt{n}} \Rightarrow \bar{y}_{i.} \pm 2.487\frac{1.7847}{\sqrt{18}} \Rightarrow$ $\bar{y}_{i.} \pm 1.05$

	Firm		
	1	2	3
Mean	9.611	9.888	7.722
95% C.I.	(8.6, 10.7)	(8.8, 10.9)	(6.7, 8.8)

There does not appear to be a significant difference between the three firms using this method of constructing C.I.'s. However, this method is very conservative. In fact, an LSD procedure would identify Firm 2 as having a significantly greater mean elapsed time than Firm 1.

9.29 a. Latin Square Design with blocking variables Farm and Fertility. The treatment is the five types of fertilizers.

b. There is significant evidence (p-value < 0.0001) the mean yields are different for the five fertilizers.

9.30 Using Tukey's W-procedure with $\alpha = 0.05, s_\epsilon^2 = MSE = .3886, q_\alpha(t, df_{error}) = q_{.05}(5, 12) = 4.52 \Rightarrow$

$W = (4.52)\sqrt{\frac{.3886}{5}} = 1.26 \Rightarrow$

	Fertilizer				
	A	B	C	D	E
Mean	5.32	6.56	7.64	7.88	8.24
Grouping	a	ab	bc	c	c

The following pairs of fertilizers have significantly different mean yields:

(A,C), (A,D), (A,E), (B,D), (B,E)