Chapter 13

Experimental Design and Analysis of Variance

Learning Objectives

1. Understand the basic principles of an experimental study.
2. Understand the difference between a completely randomized design, a randomized block design, and a factorial experiment.
3. Know the assumptions necessary to use the analysis of variance procedure.
4. Understand the use of the *F* distribution in performing the analysis of variance (ANOVA) procedure.
5. Know how to set up an ANOVA table and interpret the entries in the table.
6. Know how to use the analysis of variance procedure to determine if the means of more than two populations are equal for a completely randomized design, a randomized block design, and a factorial experiment.
7. Know how to use the analysis of variance procedure to determine if the means of more than two populations are equal for an observational study.
8. Be able to use output from computer software packages to solve experimental design problems.
9. Know how to use Fisher's least significant difference (LSD) procedure and Fisher's LSD with the Bonferroni adjustment to conduct statistical comparisons between pairs of population means.

Solutions

1. a. $\bar{\bar{x}} = (156 + 142 + 134)/3 = 144$

$$SSTR = \sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2 = 6(156 - 144)^2 + 6(142 - 144)^2 + 6(134 - 144)^2 = 1,488$$

b. MSTR = SSTR /(k – 1) = 1488/2 = 744

c. s_1^2 = 164.4 s_2^2 = 131.2 s_3^2 = 110.4

$$SSE = \sum_{j=1}^{k} (n_j - 1)s_j^2 = 5(164.4) + 5(131.2) + 5(110.4) = 2030$$

d. MSE = SSE /(n_T – k) = 2030/(18 – 3) = 135.3

e.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	*F*	*p*-value
Treatments	1,488	2	744	5.50	.0162
Error	2,030	15	135.3		
Total	3,518	17			

f. F = MSTR /MSE = 744/135.3 = 5.50

Using F table (2 degrees of freedom numerator and 15 denominator), p-value is between .01 and .025

Using Excel, the p-value corresponding to F = 5.50 is .0162.

Because p-value $\leq \alpha$ = .05, we reject the hypothesis that the means for the three treatments are equal.

2.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	*F*	*p*-value
Treatments	300	4	75	14.07	.0000
Error	160	30	5.33		
Total	460	34			

3. a. H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5$

   H_a: Not all the population means are equal

   b. Using *F* table (4 degrees of freedom numerator and 30 denominator), *p*-value is less than .01.

      Using Excel, the *p*-value corresponding to $F = 14.07$ is .0000.

      Because *p*-value $\leq \alpha = .05$, we reject H_0.

4.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	*F*	*p*-value
Treatments	150	2	75	4.80	.0233
Error	250	16	15.63		
Total	400	18			

Using *F* table (2 degrees of freedom numerator and 16 denominator), *p*-value is between .01 and .025

Using Excel, the *p*-value corresponding to $F = 4.80$ is .0233.

Because p-value $\leq \alpha = .05$, we reject the null hypothesis that the means of the three treatments are equal.

5.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Treatments	1,200	2	600	43.99	.0000
Error	600	44	13.64		
Total	1,800	46			

Using F table (2 degrees of freedom numerator and 44 denominator), p-value is less than .01

Using Excel, the p-value corresponding to $F = 43.99$ is .0000.

Because p-value $\leq \alpha = .05$, we reject the hypothesis that the treatment means are equal.

6.

	A	B	C
Sample Mean	119	107	100
Sample Variance	146.86	96.44	173.78

$$\bar{\bar{x}} = \frac{8(119) + 10(107) + 10(100)}{28} = 107.93$$

$$\text{SSTR} = \sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2 = 8(119 - 107.93)^2 + 10(107 - 107.93)^2 + 10(100 - 107.93)^2 =$$

$$1617.857$$

MSTR = SSTR /(k – 1) = 1617.857 /2 = 808.93

$$SSE = \sum_{j=1}^{k}(n_j - 1)s_j^2 = 7(146.86) + 9(96.44) + 9(173.78) = 3,460$$

MSE = SSE /(n_T – k) = 3,460 /(28 – 3) = 138.4

F = MSTR /MSE = 809.95 /138.4 = 5.85

Using F table (2 degrees of freedom numerator and 25 denominator), p-value is less than .01

Using Excel, the p-value corresponding to F = 5.85 is .0082.

Because $p\text{-value} \le \alpha = .05$, we reject the null hypothesis that the means of the three treatments are equal.

7. a.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Treatments	4,560	2	2,280	9.87	.0006
Error	6,240	27	231.11		
Total	10,800	29			

b. Using F table (2 degrees. of freedom numerator and 27 denominator), p-value is less than .01

Using Excel, the p-value corresponding to F = 9.87 is .0006.

Because $p\text{-value} \le \alpha = .05$, we reject the null hypothesis that the means of the three assembly methods are equal.

8. $\bar{\bar{x}}$ = (79 + 74 + 66)/3 = 73

$$\text{SSTR} = \sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2 = 6(79 - 73)^2 + 6(74 - 73)^2 + 6(66 - 73)^2 = 516$$

MSTR = SSTR /(k – 1) = 516/2 = 258

$s_1^2 = 34$ $s_2^2 = 20$ $s_3^2 = 32$

$$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2 = 5(34) + 5(20) + 5(32) = 430$$

MSE = SSE /(n_T – k) = 430/(18 – 3) = 28.67

F = MSTR /MSE = 258/28.67 = 9.00

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Treatments	516	2	258	9.00	.003
Error	430	15	28.67		
Total	946	17			

Using F table (2 degrees of freedom numerator and 15 denominator), p-value is less than .01.

Using Excel the p-value corresponding to F = 9.00 is .003.

Because p-value $\le \alpha = .05$, we reject the null hypothesis that the means for the three plants are equal. In other words, analysis of variance supports the conclusion that the population mean examination score at the three NCP plants are not equal.

9.

	50°	60°	70°
Sample Mean	33	29	28

Sample Variance	32	17.5	9.5

$\bar{\bar{x}}$ = (33 + 29 + 28)/3 = 30

$\text{SSTR} = \sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2 = 5(33 - 30)^2 + 5(29 - 30)^2 + 5(28 - 30)^2 = 70$

MSTR = SSTR /(k – 1) = 70 /2 = 35

$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2 = 4(32) + 4(17.5) + 4(9.5) = 236$

MSE = SSE /(n_T – k) = 236 /(15 – 3) = 19.67

F = MSTR /MSE = 35 /19.67 = 1.78

Using F table (2 degrees of freedom numerator and 12 denominator), p-value is greater than .10.

Using Excel the p-value corresponding to F = 1.78 is .2104.

Because p-value > α = .05, we cannot reject the null hypothesis that the mean yields for the three temperatures are equal.

10.

	Direct Experience	Indirect Experience	Combination
Sample Mean	17.0	20.4	25.0
Sample Variance	5.01	6.26	4.01

$\bar{\bar{x}}$ = (17 + 20.4 + 25)/3 = 20.8

$\text{SSTR} = \sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2 = 7(17 - 20.8)^2 + 7(20.4 - 20.8)^2 + 7(25 - 20.8)^2 = 225.68$

MSTR = SSTR /(k – 1) = 225.68 /2 = 112.84

$\text{SSE} = \sum_{j=1}^{k}(n_j - 1)s_j^2 = 6(5.01) + 6(6.26) + 6(4.01) = 91.68$

MSE = SSE $/(n_T - k) = 91.68/(21 - 3) = 5.09$

F = MSTR /MSE = 112.84 /5.09 = 22.17

Using F table (2 degrees of freedom numerator and 18 denominator), p-value is less than .01.

Using Excel the p-value corresponding to $F = 22.17$ is .0000.

Because $p\text{-value} \le \alpha = .05$, we reject the null hypothesis that the means for the three groups are equal.

11.

	Robot 1	Robot 2	Robot 3	Robot 4
Sample Mean	133	139	136	144
Sample Variance	47.5	50	21	54.5

$\bar{\bar{x}} = (133 + 139 + 136 + 144)/3 = 138$

$\text{SSTR} = \sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2 = 5(133 - 138)^2 + 5(139 - 138)^2 + 5(136 - 138)^2 + 5(144 - 138)^2 =$

330

MSTR = SSTR $/(k - 1) = 330/3 = 110$

$\text{SSE} = \sum_{j=1}^{k}(n_j - 1)s_j^2 = 4(47.5) + 4(50) + 4(21) + 4(54.5) = 692$

MSE = SSE $/(n_T - k) = 692/(20 - 4) = 43.25$

F = MSTR /MSE = 110 /43.25 = 2.54

Using F table (3 degrees of freedom numerator and 16 denominator), p-value is between .05 and .10.

Using Excel the p-value corresponding to $F = 2.54$ is .0931.

Because p-value $> \alpha = .05$, we cannot reject the null hypothesis that the mean drying times for the four paints are equal.

12.

	Italian	Seafood	Steakhouse
Sample Mean	17	19	24
Sample Variance	14.857	13.714	14.000

$\bar{\bar{x}} = (17 + 19 + 24)/3 = 20$

$$\text{SSTR} = \sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2 = 8(17 - 20)^2 + 8(19 - 20)^2 + 8(24 - 20)^2 = 208$$

MSTR = SSTR /$(k - 1)$ = 208/2 = 104

$$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2 = 7(14.857) + 7(13.714) + 7(14.000) = 298$$

MSE = SSE /$(n_T - k)$ = 298 /(24 – 3) = 14.19

F = MSTR /MSE = 104 /14.19 = 7.33

Using the F table (2 degrees of freedom numerator and 21 denominator), the p-value is less than .01.

Using Excel the p-value corresponding to $F = 7.33$ is .0038.

Because p-value $\leq \alpha = .05$, we reject the null hypothesis that the mean meal prices are the same for the three types of restaurants.

13. a. $\bar{\bar{x}} = (30 + 45 + 36)/3 = 37$

$$\text{SSTR} = \sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2 = 5(30 - 37)^2 + 5(45 - 37)^2 + 5(36 - 37)^2 = 570$$

MSTR = SSTR /(k – 1) = 570/2 = 285

$\text{SSE} = \sum_{j=1}^{k}(n_j - 1)s_j^2$ = 4(6) + 4(4) + 4(6.5) = 66

MSE = SSE /(n_T – k) = 66/(15 – 3) = 5.5

F = MSTR /MSE = 285/5.5 = 51.82

Using F table (2 degrees of freedom numerator and 12 denominator), p-value is less than .01.

Using Excel, the p-value corresponding to F = 51.82 is .0000.

Because p-value $\leq \alpha$ = .05, we reject the null hypothesis that the means of the three populations are equal.

b. $\text{LSD} = t_{\alpha/2}\sqrt{\text{MSE}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} = t_{.025}\sqrt{5.5\left(\frac{1}{5} + \frac{1}{5}\right)} = 2.179\sqrt{2.2} = 3.23$

$|\bar{x}_1 - \bar{x}_2| = |30 - 45| = 15 >$ LSD; significant difference

$|\bar{x}_1 - \bar{x}_3| = |30 - 36| = 6 >$ LSD; significant difference

$|\bar{x}_2 - \bar{x}_3| = |45 - 36| = 9 >$ LSD; significant difference

c. $\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2}\sqrt{\text{MSE}\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}$

$(30 - 45) \pm 2.179\sqrt{5.5\left(\frac{1}{5} + \frac{1}{5}\right)}$

–15 $\pm$ 3.23 = –18.23 to –11.77

14. a.

	Sample 1	Sample 2	Sample 3
Sample Mean	51	77	58

Sample Variance	96.67	97.34	81.99

$\bar{\bar{x}} = (51 + 77 + 58)/3 = 62$

$\text{SSTR} = \sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2 = 4(51 - 62)^2 + 4(77 - 62)^2 + 4(58 - 62)^2 = 1{,}448$

$\text{MSTR} = \text{SSTR}/(k - 1) = 1{,}448/2 = 724$

$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2 = 3(96.67) + 3(97.34) + 3(81.99) = 828$

$\text{MSE} = \text{SSE}/(n_T - k) = 828/(12 - 3) = 92$

$F = \text{MSTR}/\text{MSE} = 724/92 = 7.87$

Using F table (2 degrees of freedom numerator and 9 denominator), p-value is between .01 and .025.

Actual p-value = .0106

Because p-value $\leq \alpha = .05$, we reject the null hypothesis that the means of the three populations are equal.

b. $\text{LSD} = t_{\alpha/2}\sqrt{\text{MSE}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} = t_{.025}\sqrt{92\left(\frac{1}{4} + \frac{1}{4}\right)} = 2.262\sqrt{46} = 15.34$

$|\bar{x}_1 - \bar{x}_2| = |51 - 77| = 26 > \text{LSD}$; significant difference

$|\bar{x}_1 - \bar{x}_3| = |51 - 58| = 7 < \text{LSD}$; no significant difference

$|\bar{x}_2 - \bar{x}_3| = |77 - 58| = 19 > \text{LSD}$; significant difference

15. a.

	Manufacturer 1	Manufacturer 2	Manufacturer 3
Sample Mean	23	28	21

Sample Variance	6.67	4.67	3.33

$\bar{\bar{x}} = (23 + 28 + 21)/3 = 24$

$\text{SSTR} = \sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2 = 4(23 - 24)^2 + 4(28 - 24)^2 + 4(21 - 24)^2 = 104$

MSTR = SSTR /$(k - 1)$ = 104/2 = 52

$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2 = 3(6.67) + 3(4.67) + 3(3.33) = 44.01$

MSE = SSE /$(n_T - k)$ = 44.01/(12 – 3) = 4.89

F = MSTR /MSE = 52/4.89 = 10.63

Using F table (2 degrees of freedom numerator and 9 denominator), p-value is less than .01.

Using Excel, the p-value corresponding to F = 10.63 is .0043.

Because p-value $\leq \alpha = .05$, we reject the null hypothesis that the mean time needed to mix a batch of material is the same for each manufacturer.

b. $\text{LSD} = t_{\alpha/2}\sqrt{\text{MSE}\left(\frac{1}{n_1} + \frac{1}{n_3}\right)} = t_{.025}\sqrt{4.89\left(\frac{1}{4} + \frac{1}{4}\right)} = 2.262\sqrt{2.45} = 3.54$

Because $|\bar{x}_1 - \bar{x}_3| = |23 - 21| = 2 < 3.54,$ there does not appear to be any significant difference between the means for manufacturer 1 and manufacturer 3.

16. $\bar{x}_1 - \bar{x}_2 \pm \text{LSD}$

$23 - 28 \pm 3.54$

$-5 \pm 3.54 = -8.54$ to -1.46

17. a.

	Marketing Managers	Marketing Research	Advertising
Sample Mean	5	4.5	6
Sample Variance	.8	.3	.4

$\bar{\bar{x}} = (5 + 4.5 + 6)/3 = 5.17$

$$\text{SSTR} = \sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2 = 6(5 - 5.17)^2 + 6(4.5 - 5.17)^2 + 6(6 - 5.17)^2 = 7.00$$

MSTR = SSTR /(k – 1) = 7.00/2 = 3.5

$$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2 = 5(.8) + 5(.3) + 5(.4) = 7.50$$

MSE = SSE /(n_T – k) = 7.50/(18 – 3) = .5

F = MSTR /MSE = 3.5/.50 = 7.00

Using F table (2 degrees of freedom numerator and 15 denominator), p-value is less than .01.

Using Excel, the p-value corresponding to F = 7.00 is .0071.

Because p-value $\leq \alpha$ = .05, we reject the null hypothesis that the mean perception score is the same for the three groups of specialists.

b. Because there are only three possible pairwise comparisons we will use the Bonferroni adjustment.

α = .05/3 = .0167

Using Excel, $t_{.0167/2} = t_{.00835} = 2.694$.

$$\text{BSD} = 2.694\sqrt{\text{MSE}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} = 2.694\sqrt{.5\left(\frac{1}{6} + \frac{1}{6}\right)} = 1.0998$$

$|\bar{x}_1 - \bar{x}_2| = |5 - 4.5| = .5 < 0.956$; no significant difference

$|\bar{x}_1 - \bar{x}_3| = |5 - 6| = 1 > 0.956$; significant difference

$|\bar{x}_2 - \bar{x}_3| = |4.5 - 6| = 1.5 > 0.956$; significant difference

18. a.

	Machine 1	Machine 2	Machine 3	Machine 4
Sample Mean	7.1	9.1	9.9	11.4
Sample Variance	1.21	.93	.70	1.02

$\bar{\bar{x}} = (7.1 + 9.1 + 9.9 + 11.4)/4 = 9.38$

$\text{SSTR} = \sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2 = 6(7.1 - 9.38)^2 + 6(9.1 - 9.38)^2 + 6(9.9 - 9.38)^2 + 6(11.4 - 9.38)^2 =$

57.77

MSTR = SSTR $/(k - 1) = 57.77/3 = 19.26$

$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2 = 5(1.21) + 5(.93) + 5(.70) + 5(1.02) = 19.30$

MSE = SSE $/(n_T - k) = 19.30/(24 - 4) = .97$

F = MSTR /MSE = 19.26/.97 = 19.86

Using F table (3 degrees of freedom numerator and 20 denominator), p-value is less than .01.

Using Excel, the p-value corresponding to $F = 19.86$ is .0000.

Because p-value $\leq \alpha = .05$, we reject the null hypothesis that the mean time between breakdowns is the same for the four machines.

b. Note: $t_{\alpha/2}$ is based on 20 degrees of freedom.

$$\text{LSD} = t_{\alpha/2}\sqrt{\text{MSE}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} = t_{.025}\sqrt{0.97\left(\frac{1}{6} + \frac{1}{6}\right)} = 2.086\sqrt{.3233} = 1.19$$

$|\bar{x}_2 - \bar{x}_4| = |9.1 - 11.4| = 2.3 > \text{LSD}$; significant difference

19. C = 6 [(1,2), (1,3), (1,4), (2,3), (2,4), (3,4)]

$\alpha = .05/6 = .008$ and $\alpha/2 = .004$

Because the smallest value for $\alpha/2$ in the t table is .005, we will use $t_{.005} = 2.845$ as an approximation for $t_{.004}$ (20 degrees of freedom).

$$\text{BSD} = 2.845\sqrt{0.97\left(\frac{1}{6}+\frac{1}{6}\right)} = 1.62$$

Thus, if the absolute value of the difference between any two sample means exceeds 1.62, there is sufficient evidence to reject the hypothesis that the corresponding population means are equal.

Means	(1,2)	(1,3)	(1,4)	(2,3)	(2,4)	(3,4)
Difference	2	2.8	4.3	0.8	2.3	1.5
Significant?	Yes	Yes	Yes	No	Yes	No

20. a. Partial output is shown below:

One-way ANOVA: Attendance versus Division

Analysis of Variance for Attendance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Division	2	18109727	9054863	6.96	0.011
Error	11	14315319	1301393		
Total	13	32425045			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)

```
1140.79          55.85%       47.82%        30.33%

Means

Division          N     Mean    StDev   95% CI

North             6     7702    1301    (6677, 8727)

South             4     5566    1275    (4310, 6821)

West              4     8430    570     (7174, 9685)

Pooled StDev = 1140.79
```

Because p-value = .011 $\le \alpha$ = .05, we reject the null hypothesis that the mean attendance values are equal.

b. $n_1 = 6$ $n_2 = 4$ $n_3 = 4$

$t_{\alpha/2}$ is based upon 11 degrees of freedom

Comparing North and South

$$\text{LSD} = t_{.025}\sqrt{1,301,393\left(\frac{1}{6}+\frac{1}{4}\right)} = 2.201\sqrt{1,301,393\left(\frac{1}{6}+\frac{1}{4}\right)} = 1620.76$$

$|7702-5566| = 2136 > \text{LSD}$; significant difference

Comparing North and West

$$\text{LSD} = t_{.025}\sqrt{1,301,393\left(\frac{1}{6}+\frac{1}{4}\right)} = 2.201\sqrt{1,301,393\left(\frac{1}{6}+\frac{1}{4}\right)} = 1620.76$$

$|7702-8430| = 728 < \text{LSD}$; no significant difference

Comparing South and West

$$\text{LSD} = t_{.025}\sqrt{1,301,393\left(\frac{1}{4}+\frac{1}{4}\right)} = 2.201\sqrt{1,301,393\left(\frac{1}{4}+\frac{1}{4}\right)} = 1775.45$$

$|5566 - 8430| = 2864 >$ LSD; significant difference

The difference in the mean attendance among the three divisions is the result of low attendance in the South division.

21. Treatment Means

$\bar{x}_{.1} = 13.6$ $\bar{x}_{.2} = 11.0$ $\bar{x}_{.3} = 10.6$

Block Means:

$\bar{x}_{1.} = 9$ $\bar{x}_{2.} = 7.67$ $\bar{x}_{3.} = 15.67$ $\bar{x}_{4.} = 18.67$ $\bar{x}_{5.} = 7.67$

Overall Mean:

$\bar{\bar{x}} = 176/15 = 11.73$

Step 1

$$\text{SST} = \sum_i \sum_j \left(x_{ij} - \bar{\bar{x}}\right)^2 = (10 - 11.73)^2 + (9 - 11.73)^2 + \cdot\cdot\cdot + (8 - 11.73)^2 = 354.93$$

Step 2

$$\text{SSTR} = b\sum_j \left(\bar{x}_{.j} - \bar{\bar{x}}\right)^2 = 5\ [\ (13.6 - 11.73)^2 + (11.0 - 11.73)^2 + (10.6 - 11.73)^2\ ] = 26.53$$

Step 3

$$\text{SSBL} = k\sum_i \left(\bar{x}_{i.} - \bar{\bar{x}}\right)^2 = 3\ [\ (9 - 11.73)^2 + (7.67 - 11.73)^2 + (15.67 - 11.73)^2 + (18.67 - 11.73)^2 + (7.67 - 11.73)^2\ ] = 312.32$$

Step 4

SSE = SST – SSTR – SSBL = 354.93 – 26.53 – 312.32 = 16.08

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	*F*	*p*-value
Treatments	26.53	2	13.27	6.60	.0203
Blocks	312.32	4	78.08		
Error	16.08	8	2.01		
Total	354.93	14			

Using *F* table (2 degrees of freedom numerator and 8 denominator), *p*-value is between .01 and .025.

Using Excel, the *p*-value corresponding to $F = 6.60$ is .0203.

Because *p*-value $\leq \alpha = .05$, we reject the null hypothesis that the means of the three treatments are equal.

22.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	*F*	*p*-value
Treatments	310	4	77.5	17.69	.0005
Blocks	85	2	42.5		
Error	35	8	4.38		
Total	430	14			

Using *F* table (4 degrees of freedom numerator and 8 denominator), *p*-value is less than

.01.

Using Excel, the *p*-value corresponding to $F = 17.69$ is .0005.

Because *p*-value $\leq \alpha = .05$, we reject the null hypothesis that the means of the treatments are equal.

23.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	*F*	*p*-value
Treatments	900	3	300	12.60	.0001
Blocks	400	7	57.14		
Error	500	21	23.81		
Total	1,800	31			

Using *F* table (3 degrees of freedom numerator and 21 denominator), *p*-value is less than .01.

Using Excel, the *p*-value corresponding to $F = 12.60$ is .0001.

Because *p*-value$\leq \alpha = .05$, we reject the null hypothesis that the means of the treatments are equal.

24. Treatment Means:

$\bar{x}_{\cdot 1} = 56$ $\bar{x}_{\cdot 2} = 44$

Block Means:

$\bar{x}_{1\cdot} = 46$ $\bar{x}_{2\cdot} = 49.5$ $\bar{x}_{3\cdot} = 54.5$

Overall Mean:

$\bar{\bar{x}} = 300/6 = 50$

Step 1

$$SST = \sum_i \sum_j \left(x_{ij} - \bar{\bar{x}}\right)^2 = (50 - 50)^2 + (42 - 50)^2 + \cdots + (46 - 50)^2 = 310$$

Step 2

$$SSTR = b\sum_j \left(\bar{x}_{.j} - \bar{\bar{x}}\right)^2 = 3\,[\,(56 - 50)^2 + (44 - 50)^2\,] = 216$$

Step 3

$$SSBL = k\sum_i \left(\bar{x}_{i.} - \bar{\bar{x}}\right)^2 = 2\,[\,(46 - 50)^2 + (49.5 - 50)^2 + (54.5 - 50)^2\,] = 73$$

Step 4

SSE = SST – SSTR – SSBL = 310 – 216 – 73 = 21

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	*F*	*p*-value
Treatments	216	1	216	20.57	.0453
Blocks	73	2	36.5		
Error	21	2	10.5		
Total	310	5			

Using *F* table (1 degree of freedom numerator and 2 denominator), *p*-value is between .025 and .05.

Using Excel, the *p*-value corresponding to $F = 20.57$ is .0453.

Because p-value $\leq \alpha$ = .05, we reject the null hypothesis that the mean tune-up times are the same for both analyzers.

25. The blocks correspond to the four trips and the treatments correspond to the three travel agency websites.

Partial Excel two-way ANOVA output follows.

ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Trip	25487.68	3	8495.894	578.7412	8.92E-08	4.757062663
Website	75.78302	2	37.89151	2.581174	0.155306	5.14325285
Error	88.07972	6	14.67995			
Total	25651.55	11				

Because the p-value for Website (.1553) is greater than α = .05, we cannot reject the null hypothesis that there is no difference in price among the three websites.

26. a. Treatment Means:

$\bar{x}_{.1} = 502$ $\bar{x}_{.2} = 515$ $\bar{x}_{.3} = 494$

Block Means:

$\bar{x}_{1.} = 530$ $\bar{x}_{2.} = 590$ $\bar{x}_{3.} = 458$ $\bar{x}_{4.} = 560$ $\bar{x}_{5.} = 448$ $\bar{x}_{6.} = 436$

Overall Mean:

$\bar{\bar{x}} = 9066/18 = 503.67$

Step 1

$$\text{SST} = \sum_i \sum_j \left(x_{ij} - \bar{\bar{x}}\right)^2 = (526 - 503.67)^2 + (534 - 503.67)^2 + \cdot\cdot\cdot + (420 - 503.67)^2 =$$

65,798

Step 2

$$\text{SSTR} = b\sum_j \left(\bar{x}_{.j} - \bar{\bar{x}}\right)^2 = 6[\ (502 - 503.67)^2 + (515 - 503.67)^2 + (494 - 503.67)^2\ ] = 1348$$

Step 3

$$\text{SSBL} = k\sum_i \left(\bar{x}_{i.} - \bar{\bar{x}}\right)^2 = 3\ [\ (530 - 503.67)^2 + (590 - 503.67)^2 + \cdot\cdot\cdot + (436 - 503.67)^2\ ] =$$

63,250

Step 4

SSE = SST – SSTR – SSBL = 65,798 – 1348 – 63,250= 1200

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	*F*	*p*-value
Treatments	1348	2	674	5.62	.0231
Blocks	63,250	5	12,650		
Error	1200	10	120		
Total	65,798	17			

Using *F* table (2 degrees of freedom numerator and 10 denominator), *p*-value is between .01 and .025.

Using Excel, the *p*-value corresponding to $F = 5.62$ is .0231.

Because *p*-value $\leq \alpha = .05$, we reject the null hypothesis that the mean scores for the three parts of the SAT are equal.

b. The mean test scores for the three sections are 502 for critical reading, 515 for mathematics, and 494 for writing. Because the writing section has the lowest average score, this section appears to give the students the most trouble.

27. Partial Excel output is shown as follows.

ANOVA						
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Rows	931.82	9	103.5356	6.7342	0.0003	2.4563
Columns	37.20267	2	18.6013	1.2099	0.3214	3.5546
Error	276.744	18	15.374667			
Total	1245.767	29				

The p-value corresponding to Design (Columns) is .3214; because the p-value = .3214 > .05, we cannot reject the null hypothesis that the designs are equally preferred.

28.

		Factor B Level 1	Factor B Level 2	Factor B Level 3	Factor A Means
Factor A	Level 1	= 150	= 78	= 84	= 104
Factor A	Level 2	= 110	= 116	= 128	= 118
Factor B	Means	= 130	= 97	= 106	= 111

Step 1

$$\text{SST} = \sum_i \sum_j \sum_k \left(x_{ijk} - \bar{\bar{x}}\right)^2 = (135 - 111)^2 + (165 - 111)^2 + \cdots + (136 - 111)^2 = 9{,}028$$

Step 2

$$\text{SSA} = br \sum_i \left(\bar{x}_{i\cdot} - \bar{\bar{x}}\right)^2 = 3\,(2)\,[\,(104 - 111)^2 + (118 - 111)^2\,] = 588$$

Step 3

$$SSB = ar\sum_j (\bar{x}_{.j} - \bar{\bar{x}})^2 = 2\ (2)\ [\ (130 - 111)^2 + (97 - 111)^2 + (106 - 111)^2\ ] = 2{,}328$$

Step 4

$$SSAB = r\sum_i \sum_j (\bar{x}_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{\bar{x}})^2 = 2\ [\ (150 - 104 - 130 + 111)^2 + (78 - 104 - 97 + 111)^2 + \cdot$$

$$\cdot + (128 - 118 - 106 + 111)^2\ ] = 4{,}392$$

Step 5

SSE = SST – SSA – SSB – SSAB = 9,028 – 588 – 2,328 – 4,392 = 1,720

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	*F*	*p*-value
Factor A	588	1	588	2.05	.2022
Factor B	2328	2	1164	4.06	.0767
Interaction	4392	2	2196	7.66	.0223
Error	1720	6	286.67		
Total	9028	11			

Factor A: $F = 2.05$

Using *F* table (1 degree of freedom numerator and 6 denominator), *p*-value is greater than .10.

Using Excel, the *p*-value corresponding to $F = 2.05$ is .2022.

Because *p*-value $> \alpha = .05$, Factor A is not significant.

Factor B: $F = 4.06$

Using F table (2 degrees of freedom numerator and 6 denominator), p-value is between .05 and .10.

Using Excel, the p-value corresponding to $F = 4.06$ is .0767.

Because p-value $> \alpha = .05$, Factor B is not significant.

Interaction: $F = 7.66$

Using F table (2 degrees of freedom numerator and 6 denominator), p-value is between .01 and .025.

Using Excel, the p-value corresponding to $F = 7.66$ is .0223.

Because p-value $\leq \alpha = .05$, Interaction is significant.

29.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Factor A	26	3	8.67	3.72	.0250
Factor B	23	2	11.50	4.94	.0160
Interaction	175	6	29.17	12.52	.0000
Error	56	24	2.33		
Total	280	35			

Using F table for Factor A (3 degrees of freedom numerator and 24 denominator), p-value is .025.

Because p-value $\leq \alpha = .05$, Factor A is significant.

Using F table for Factor B (2 degrees of freedom numerator and 24 denominator), p-

value is between .01 and .025.

Using Excel, the p-value corresponding to $F = 4.94$ is .0160.

Because $p\text{-value} \leq \alpha = .05$, Factor B is significant.

Using F table for Interaction (6 degrees of freedom numerator and 24 denominator), p-value is less than .01.

Using Excel, the p-value corresponding to $F = 12.52$ is .0000.

Because $p\text{-value} \leq \alpha = .05$, Interaction is significant.

30. Factor A is navigation menu position; Factor B is amount of text entry required.

		Factor B		Factor A
		Small	Large	Means
	A	$\bar{x}_{11} = 10$	$\bar{x}_{12} = 10$	$\bar{x}_{1.} = 10$
Factor A	B	$\bar{x}_{21} = 18$	$\bar{x}_{22} = 28$	$\bar{x}_{2.} = 23$
	C	$\bar{x}_{31} = 14$	$\bar{x}_{32} = 16$	$\bar{x}_{3.} = 15$
Factor B	Means	$\bar{x}_{.1} = 14$	$\bar{x}_{.2} = 18$	$\bar{\bar{x}} = 16$

Step 1

$$\text{SST} = \sum_i \sum_j \sum_k \left(x_{ijk} - \bar{\bar{x}}\right)^2 = (8 - 16)^2 + (12 - 16)^2 + (12 - 16)^2 + \cdot \cdot \cdot + (14 - 16)^2 = 664$$

Step 2

$$\text{SSA} = br \sum_i \left(\bar{x}_{i.} - \bar{\bar{x}}\right)^2 = 2\,(2)\,[\,(10 - 16)^2 + (23 - 16)^2 + (15 - 16)^2\,] = 344$$

Step 3

$\text{SSB} = ar\sum_j \left(\bar{x}_{.j} - \bar{\bar{x}}\right)^2 = 3\ (2)\ [\ (14 - 16)^2 + (18 - 16)^2\ ] = 48$

Step 4

$\text{SSAB} = r\sum_i \sum_j \left(\bar{x}_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{\bar{x}}\right)^2 = 2\ [\ (10 - 10 - 14 + 16)^2 + \cdot \cdot \cdot + (16 - 15 - 18 + 16)^2\ ] =$

56

Step 5

SSE = SST – SSA – SSB – SSAB = 664 – 344 – 48 – 56 = 216

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	*F*	*p*-value
Factor A	344	2	172	172/36 = 4.78	.0574
Factor B	48	1	48	48/36 = 1.33	.2921
Interaction	56	2	28	28/36 = 0.78	.5008
Error	216	6	36		
Total	664	11			

Using *F* table for Factor A (2 degrees of freedom numerator and 6 denominator), *p*-value is between .05 and .10.

Using Excel the *p*-value corresponding to $F = 4.78$ is .0574.

Because *p*-value > $\alpha = .05$, Factor A is not significant; there is not sufficient evident to suggest a difference due to the navigation menu position.

Using *F* table for Factor B (1 degree of freedom numerator and 6 denominator), *p*-value is greater than .10.

Using Excel the *p*-value corresponding to F =1.33 is .2921.

Because *p*-value $> \alpha$ = .05, Factor B is not significant; there is not a significant difference due to amount of required text entry.

Using F table for Interaction (2 degrees of freedom numerator and 6 denominator), *p*-value is greater than .10.

Using Excel, the *p*-value corresponding to F = 0.78 is .5008.

Because *p*-value $> \alpha$ = .05, Interaction is not significant.

31. Factor A is method of loading and unloading; Factor B is type of ride.

		Factor B			Factor A
		Roller Coaster	Screaming Demon	Log Flume	Means
Factor A	Method 1	$\bar{x}_{11} = 42$	$\bar{x}_{12} = 48$	$\bar{x}_{13} = 48$	$\bar{x}_{1.} = 46$
	Method 2	$\bar{x}_{21} = 50$	$\bar{x}_{22} = 48$	$\bar{x}_{23} = 46$	$\bar{x}_{2.} = 48$
Factor B	Means	$\bar{x}_{.1} = 46$	$\bar{x}_{.2} = 48$	$\bar{x}_{.3} = 47$	$\bar{\bar{x}} = 47$

Step 1

$$\text{SST} = \sum_i \sum_j \sum_k \left(x_{ijk} - \bar{\bar{x}}\right)^2 = (41 - 47)^2 + (43 - 47)^2 + \cdots + (44 - 47)^2 = 136$$

Step 2

$$\text{SSA} = br \sum_i \left(\bar{x}_{i.} - \bar{\bar{x}}\right)^2 = 3\ (2)\ [\ (46 - 47)^2 + (48 - 47)^2\ ] = 12$$

Step 3

$$\text{SSB} = ar\sum_j \left(\bar{x}_{.j} - \bar{\bar{x}}\right)^2 = 2\,(2)\,[\,(46-47)^2 + (48-47)^2 + (47-47)^2\,] = 8$$

Step 4

$$\text{SSAB} = r\sum_i \sum_j \left(\bar{x}_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{\bar{x}}\right)^2 = 2\,[\,(41-46-46+47)^2 + \cdot\cdot\cdot + (44-48-47+47)^2\,] =$$

56

Step 5

SSE = SST – SSA – SSB – SSAB = 136 – 12 – 8 – 56 = 60

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	*F*	*p*-value
Factor A	12	1	12	12/10 = 1.2	.3153
Factor B	8	2	4	4/10 = .4	.6870
Interaction	56	2	28	28/10 = 2.8	.1384
Error	60	6	10		
Total	136	11			

Using *F* table for Factor A (1 degree of freedom numerator and 6 denominator), *p*-value is greater than .10

Using Excel, the *p*-value corresponding to $F = 1.2$ is .3153.

Because *p*-value $> \alpha = .05$, Factor A is not significant.

Using *F* table for Factor B (2 degrees of freedom numerator and 6 denominator), *p*-value is greater than .10

Using Excel, the p-value corresponding to $F = .4$ is .6870.

Because $p\text{-value} > \alpha = .05$, Factor B is not significant.

Using F table for Interaction (2 degrees of freedom numerator and 6 denominator), p-value is greater than .10

Using Excel, the p-value corresponding to $F = 2.8$ is .1384.

Because $p\text{-value} > \alpha = .05$, Interaction is not significant.

32. Factor A is class of vehicle tested (small car, midsize car, small SUV, and midsize SUV) and Factor B is Type (hybrid or conventional). The data in tabular format follow.

	Hybrid	Conventional
Small Car	15.7	11.8
	18.6	13.5
Midsize Car	11.4	9.7
	13.5	10.6
Small SUV	11.4	8.9
	11.8	9.3
Midsize SUV	9.7	8.0
	10.2	7.6

Summary statistics for the preceding data follow.

	Hybrid	Conventional	
Small Car	$\bar{x}_{11} = 17.15$	$\bar{x}_{12} = 12.65$	$\bar{x}_{1.} = 14.90$

Midsize Car	$\bar{x}_{21} = 12.45$	$\bar{x}_{22} = 10.15$	$\bar{x}_{2.} = 11.30$
Small SUV	$\bar{x}_{31} = 11.60$	$\bar{x}_{32} = 9.10$	$\bar{x}_{3.} = 10.35$
Midsize SUV	$\bar{x}_{41} = 9.95$	$\bar{x}_{42} = 7.80$	$\bar{x}_{4.} = 8.88$
	$\bar{x}_{.1} = 12.79$	$\bar{x}_{.2} = 9.93$	$\bar{\bar{x}} = 11.36$

Step 1

$$\text{SST} = \sum_i \sum_j \sum_k \left(x_{ijk} - \bar{\bar{x}}\right)^2 = (15.7 - 11.36)^2 + (18.6 - 11.36)^2 + \cdots + (7.6 - 11.36)^2 = 123.96$$

Step 2

$$\text{SSA} = br\sum_i \left(\bar{x}_{i.} - \bar{\bar{x}}\right)^2 = 2(2)\,[(14.90 - 11.36)^2 + (11.30 - 11.36)^2 + (10.35 - 11.36)^2 + (8.88 - 11.36)^2] = 78.82$$

Step 3

$$\text{SSB} = ar\sum_j \left(\bar{x}_{.j} - \bar{\bar{x}}\right)^2 = 4(2)\,[(12.79 - 11.36)^2 + (9.93 - 11.36)^2] = 32.72$$

Step 4

$$\text{SSAB} = r\sum_i \sum_j \left(\bar{x}_{ij} - \bar{x}_{i.} - \bar{x}_{.j} + \bar{\bar{x}}\right)^2 = 2[(17.15 - 14.90 - 12.79 + 11.36)^2 + (12.65 - 14.90 - 9.93 + 11.36)^2 + \cdots + (7.80 - 8.88 - 9.93 + 11.36)^2] = 3.64$$

Step 5

$$\text{SSE} = \text{SST} - \text{SSA} - \text{SSB} - \text{SSAB} = 123.96 - 78.82 - 32.72 - 3.64 = 8.78$$

The Excel readout for ANOVA with Replication is reproduced below. The slightly different numbers arise from the 2 decimal places rounding in the preceding manual calculations.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	*F*	*p*-value
Factor A	78.92	3	26.307	24.40	.0002
Factor B	32.78	1	32.776	30.40	.0006
Interaction	3.64	3	1.212	1.12	.3952
Error	8.63	8	1.078		
Total	123.96	15			

Conclusions

Factor A: Because p-value = .0002 < α = .05, Factor A (Class) is significant.

Factor B: Because p-value = .0006 < α = .05, Factor B (Type) is significant.

Interaction: Because p-value = .3952 > α = .05, Interaction is not significant.

The class of vehicles has a significant effect on kilometers per liter with cars showing more kilometers per liter than SUVs. The type of vehicle also has a significant effect, with hybrids having more kilometers per liter than conventional vehicles. There is no evidence of a significant interaction effect.

33. Factor A is time pressure (low and moderate); Factor B is level of knowledge (naïve, declarative and procedural).

= (1.13 + 1.56 + 2.00)/3 = 1.563

$\bar{x}_{2\cdot} = (0.48 + 1.68 + 2.86)/3 = 1.673$

$\bar{x}_{\cdot 1} = (1.13 + 0.48)/2 = 0.805$

$\bar{x}_{\cdot 2} = (1.56 + 1.68)/2 = 1.620$

$\bar{x}_{\cdot 3} = (2.00 + 2.86)/2 = 2.43$

$\bar{\bar{x}} = (1.13 + 1.56 + 2.00 + 0.48 + 1.68 + 2.86)/6 = 1.618$

Step 1

SST = 327.50 (given in problem statement)

Step 2

$$\text{SSA} = br\sum_i \left(\bar{x}_{i\cdot} - \bar{\bar{x}}\right)^2 = 3(25)[(1.563 - 1.618)^2 + (1.673 - 1.618)^2] = 0.4538$$

Step 3

$$\text{SSB} = ar\sum_j \left(\bar{x}_{\cdot j} - \bar{\bar{x}}\right)^2 = 2(25)[(0.805 - 1.618)^2 + (1.62 - 1.618)^2 + (2.43 - 1.618)^2] = 66.0159$$

Step 4

$$\text{SSAB} = r\sum_i \sum_j \left(\bar{x}_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \bar{\bar{x}}\right)^2 = 25[(1.13 - 1.563 - 0.805 + 1.618)^2 + (1.56 - 1.563 - 1.62 + 1.618)^2 + \cdots + (2.86 - 1.673 - 2.43 + 1.618)^2] = 14.2525$$

Step 5

SSE = SST – SSA – SSB – SSAB = 327.50 – 0.4538 – 66.0159 – 14.2525

Source of	Sum of	Degrees of	Mean	*F*	*p*-value

Variation	Squares	Freedom	Square		
Factor A	0.4538	1	0.4538	0.2648	.6076
Factor B	66.0159	2	33.0080	19.2608	.0000
Interaction	14.2525	2	7.1263	4.1583	.0176
Error	246.7778	144	1.7137		
Total	327.5000	149			

Factor A: Using Excel, the p-value corresponding to F = .2648 is .6076. Because p-value $> \alpha$ = .05, Factor A (time pressure) is not significant.

Factor B: Using Excel, the p-value corresponding to F = 19.2608 is .0000. Because p-value $\leq \alpha$ = .05, Factor B (level of knowledge) is significant.

Interaction: Using Excel, the p-value corresponding to F = 4.1583 is .0176. Because p-value $\leq \alpha$ = .05, Interaction is significant.

34.

	x	y	z
Sample Mean	92	97	84
Sample Variance	30	6	35.33

$\bar{\bar{x}} = (92 + 97 + 44)/3 = 91$

$\text{SSTR} = \sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2 = 4(92 - 91)^2 + 4(97 - 91)^2 + 4(84 - 91)^2 = 344$

MSTR = SSTR /(k – 1) = 344 /2 = 172

$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2 = 3(30) + 3(6) + 3(35.33) = 213.99$

MSE = SSE /($n_T - k$) = 213.99 /(12 – 3) = 23.78

F = MSTR /MSE = 172 /23.78 = 7.23

Using F table (2 degrees of freedom numerator and 9 denominator), p-value is between .01 and .025

Using Excel, the p-value corresponding to F = 7.23 is .0134.

Because p-value $\le \alpha$ = .05, we reject the null hypothesis that the mean absorbency ratings for the three brands are equal.

35.

	Lawyer	Physical Therapist	Cabinet Maker	Systems Analyst
Sample Mean	50.0	63.7	69.1	61.2
Sample Variance	124.22	164.68	105.88	136.62

$$\bar{\bar{x}} = \frac{50.0 + 63.7 + 69.1 + 61.2}{4} = 61$$

$$\text{SSTR} = \sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2 = 10(50.0 - 61)^2 + 10(63.7 - 61)^2 + 10(69.1 - 61)^2 + 10(61.2 - 61)^2$$

$$= 1939.4$$

MSTR = SSTR /(k – 1) = 1939.4 /3 = 646.47

$$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2 = 9(124.22) + 9(164.68) + 9(105.88) + 9(136.62) = 4{,}782.60$$

MSE = SSE /($n_T - k$) = 4782.6 /(40 – 4) = 132.85

F = MSTR /MSE = 646.47 /132.85 = 4.87

Using F table (3 degrees of freedom numerator and 36 denominator), p-value is less than .01.

Using Excel, the p-value corresponding to $F = 4.87$ is .0061.

Because p-value $\leq \alpha = .05$, we reject the null hypothesis that the mean job satisfaction rating is the same for the four professions.

36. The blocks correspond to the 10 dates on which the data were collected (Date) and the treatments correspond to the four cities (City).

Partial ANOVA output follows.

```
Analysis of Variance for Ozone Level

Source          DF          SS          MS          F          P

Date            9           903.02      100.34      4.55       0.001

City            3           160.08      53.36       2.42       0.088

Error           27          595.68      22.06

Total           39          1658.78

S = 4.69702     R-Sq = 64.09%           R-Sq(adj) = 48.13%
```

Because the p-value for City (.088) is greater than $\alpha = .05$, there is no significant difference in the mean ozone level among the four cities. But, if the level of significance was $\alpha = .10$, the difference would have been significant.

37. Partial output is shown below:

```
One-way ANOVA: Northeast, Midwest, South, West

      Method

      Null hypothesis             All means are equal
```

Alternative hypothesis At least one mean is different

Significance level $\alpha = 0.05$

Equal variances were assumed for the analysis.

Factor Information

Factor	Levels	Values
Factor	4	Northeast, Midwest, South, West

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Factor	3	1758	586.1	3.13	0.029
Error	94	17588	187.1		
Total	97	19346			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
13.6786	9.09%	6.19%	1.36%

Means

Factor	N	Mean	StDev	95% CI
Northeast	11	38.54	13.00	(30.35, 46.73)
Midwest	30	41.84	16.61	(36.88, 46.80)
South	32	35.13	12.37	(30.33, 39.94)
West	25	30.85	11.47	(25.42, 36.28)

Pooled StDev = 13.6786

Because the *p*-value = .029 is less than $\alpha = .05$, we reject the null hypothesis that the percentage of 17- to 24-year-olds who are attending college is the same for the four geographic regions. The percentage of 17- to 24-year-olds who are attending college was highest in the Midwest (41.84%).

38.

	Method A	Method B	Method C
Sample Mean	90	84	81
Sample Variance	98.00	168.44	159.78

$\bar{\bar{x}} = (90 + 84 + 81)/3 = 85$

$\text{SSTR} = \sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2 = 10(90 - 85)^2 + 10(84 - 85)^2 + 10(81 - 85)^2 = 420$

$\text{MSTR} = \text{SSTR}/(k - 1) = 420/2 = 210$

$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2 = 9(98.00) + 9(168.44) + 9(159.78) = 3{,}836$

$\text{MSE} = \text{SSE}/(n_T - k) = 3{,}836/(30 - 3) = 142.07$

$F = \text{MSTR}/\text{MSE} = 210/142.07 = 1.48$

Using F table (2 degrees of freedom numerator and 27 denominator), p-value is greater than .10.

Using Excel, the p-value corresponding to $F = 1.48$ is .2455.

Because p-value $> \alpha = .05$, we can not reject the null hypothesis that the means are equal.

39. a.

	Nurse	Tax Auditor	Fast-Food Worker
Sample Mean	4.25	5.25	5.75
Sample Variance	1.07	1.07	1.36

$\bar{\bar{x}} = (4.25 + 5.25 + 5.75) / 3 = 5.08$

$\text{SSTR} = \sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2 = 8(4.25 - 5.08)^2 + 8(5.25 - 5.08)^2 + 8(5.75 - 5.08)^2 = 9.33$

$\text{MSTR} = \text{SSTR} /(k - 1) = 9.33 /2 = 4.67$

$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2 = 7(1.07) + 7(1.07) + 7(1.36) = 24.5$

$\text{MSE} = \text{SSE} /(n_T - k) = 24.5 /(24 - 3) = 1.17$

$F = \text{MSTR} /\text{MSE} = 4.67 /1.17 = 3.99$

Using F table (2 degrees of freedom numerator and 21 denominator), p-value is between .025 and .05.

Using Excel, the p-value corresponding to $F = 3.99$ is .0340.

Because $p\text{-value} \leq \alpha = .05$, we reject the null hypothesis that the mean scores are the same for the three professions.

b. $\text{LSD} = t_{\alpha/2}\sqrt{\text{MSE}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)} = 2.080\sqrt{1.17\left(\frac{1}{8} + \frac{1}{8}\right)} = 1.12$

Because the absolute value of the difference between the sample means for nurses and tax auditors is $|4.25 - 5.25| = 1$, we cannot reject the null hypothesis that the two population means are equal.

40. a. Treatment Means:

$\bar{x}_{.1} = 9.66$ $\bar{x}_{.2} = 10.50$ $\bar{x}_{.3} = 10.94$

Block Means:

$\bar{x}_{1\cdot} = 8.33$ $\bar{x}_{2\cdot} = 10.87$ $\bar{x}_{3\cdot} = 13.13$ $\bar{x}_{4\cdot} = 10.03$ $\bar{x}_{5\cdot} = 9.47$

Overall Mean:

$\bar{\bar{x}} = 155.5 / 15 = 10.37$

Step 1

$$\text{SST} = \sum_i \sum_j \left(x_{ij} - \bar{\bar{x}}\right)^2 = (7.6 - 10.37)^2 + (8.9 - 10.37)^2 + \cdots + (10.2 - 10.37)^2 = 45.41$$

Step 2

$$\text{SSTR} = b\sum_j \left(\bar{x}_{.j} - \bar{\bar{x}}\right)^2 = 5\ [\ (9.66 - 10.37)^2 + (10.50 - 10.37)^2 + (10.94 - 10.37)^2] = 4.23$$

Step 3

$$\text{SSBL} = k\sum_i \left(\bar{x}_{i\cdot} - \bar{\bar{x}}\right)^2 = 3\ [\ (8.33 - 10.37)^2 + (10.87 - 10.37)^2 + \cdots + (9.47 - 10.37)^2] = 38.86$$

Step 4

SSE = SST – SSTR – SSBL = 45.41 – 4.23 – 38.86 = 2.32

The Excel readout for ANOVA with Replication is reproduced below. The slightly different numbers arise from the 2 decimal places rounding in the preceding manual calculations.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Treatment	4.23	2	2.11	7.34	.0155
Blocks	38.88	4	9.72	33.75	
Error	2.30	8	0.29		
Total	45.41	14			

Using F table (2 degrees of freedom numerator and 8 denominator), p-value is between .01 and .025.

Using Excel, the p-value corresponding to $F = 7.34$ is .0155.

Because $p\text{-value} \le \alpha = .05$, we reject the null hypothesis that the mean kilometers per liter ratings for the three brands of gasoline are equal.

b.

	I	II	III
Sample Mean	9.66	10.50	10.94
Sample Variance	3.81	1.68	4.81

$\bar{\bar{x}} = (9.66 + 10.50 + 10.94)/3 = 10.37$

$\text{SSTR} = \sum_{j=1}^{k} n_j \left(\bar{x}_j - \bar{\bar{x}}\right)^2 = 5(9.66 - 10.37)^2 + 5(10.50 - 10.37)^2 + 5(10.94 - 10.37)^2 = 4.23$

$\text{MSTR} = \text{SSTR}/(k - 1) = 4.23/2 = 2.115$

$\text{SSE} = \sum_{j=1}^{k} (n_j - 1)s_j^2 = 4(3.81) + 4(1.68) + 4(4.81) = 41.2$

$\text{MSE} = \text{SSE}/(n_T - k) = 41.2/(15 - 3) = 3.43$

F = MSTR /MSE = 2.115 /3.43 = .62

Using F table (2 degrees of freedom numerator and 12 denominator), p-value is greater than .10.

Using Excel, the p-value corresponding to F = .62 is .55

Because p-value $> \alpha$ = .05, we cannot reject the null hypothesis that the mean kilometers per liter ratings for the three brands of gasoline are equal.

Thus, we must remove the block effect in order to detect a significant difference due to the brand of gasoline. The following table illustrates the relationship between the randomized block design and the completely randomized design.

Sum of Squares	Randomized Block Design	Completely Randomized Design
SST	45.41	45.41
SSTR	4.23	4.23
SSBL	38.86	does not exist
SSE	2.32	41.2

Note that SSE for the completely randomized design is the sum of SSBL (38.86) and SSE (2.32) for the randomized block design. This illustrates that the effect of blocking is to remove the block effect from the error sum of squares; thus, the estimate of σ^2 for the randomized block design is substantially smaller than it is for the completely randomized design.

41. Partial Excel output follows.

ANOVA

Source of Variation	SS	df	MS	F	*P*-value	*F* crit
Rows	0.2349	9	0.0261	0.4017	0.9180	2.4563
Columns	4.4855	2	2.2428	34.5144	6.93E-07	3.5546
Error	1.1697	18	0.0650			
Total	5.8901	29				

The label Rows corresponds to the blocks in the problem (week), and the label Columns corresponds to the treatments (Show).

Because the *p*-value corresponding to Columns is less than $\alpha = .05$, there is a significant difference in the mean viewing audience for the three late night talk shows.

42. The blocks correspond to the 12 golfers (Golfer) and the treatments correspond to the three designs (Design).

The ANOVA output follows.

```
ANOVA: Distance versus Design, Golfer

      Analysis of Variance for Distance

      Source          DF     SS          MS         F           P

      Design          2      3032.0      1516.0     12.89       0.000

      Golfer          11     5003.3      454.8      3.87        0.003

      Error           22     2586.7      117.6

      Total           35     10622.0

      S = 10.8432      R-Sq = 75.65%     R-Sq(adj) = 61.26%
```

Because the *p*-value for Design (.000) is less than $\alpha = .05$, there is a significant difference in the mean driving distance for the three designs.

43.

		Factor B			Factor A
		Spanish	French	German	Means
Factor A	System 1	$\bar{x}_{11} = 10$	$\bar{x}_{12} = 12$	$\bar{x}_{13} = 14$	$\bar{x}_{1\cdot} = 12$
	System 2	$\bar{x}_{21} = 8$	$\bar{x}_{22} = 15$	$\bar{x}_{23} = 19$	$\bar{x}_{2\cdot} = 14$
Factor B	Means	$\bar{x}_{\cdot 1} = 9$	$\bar{x}_{\cdot 2} = 13.5$	$\bar{x}_{\cdot 3} = 16.5$	$\bar{\bar{x}} = 13$

Step 1

$$\text{SST} = \sum_i \sum_j \sum_k \left(x_{ijk} - \bar{\bar{x}}\right)^2 = (8 - 13)^2 + (12 - 13)^2 + \cdots + (22 - 13)^2 = 204$$

Step 2

$$\text{SSA} = br\sum_i \left(\bar{x}_{i\cdot} - \bar{\bar{x}}\right)^2 = 3\ (2)\ [\ (12 - 13)^2 + (14 - 13)^2\ ] = 12$$

Step 3

$$\text{SSB} = ar\sum_j \left(\bar{x}_{\cdot j} - \bar{\bar{x}}\right)^2 = 2\ (2)\ [\ (9 - 13)^2 + (13.5 - 13)^2 + (16.5 - 13)^2\ ] = 114$$

Step 4

$$\text{SSAB} = r\sum_i \sum_j \left(\bar{x}_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \bar{\bar{x}}\right)^2 = 2\ [(8 - 12 - 9 + 13)^2 + \cdots + (22 - 14 - 16.5 + 13)^2] = 26$$

Step 5

SSE = SST – SSA – SSB – SSAB = 204 – 12 – 114 – 26 = 52

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Factor A	12	1	12	1.38	.2846
Factor B	114	2	57	6.57	.0308
Interaction	26	2	12	1.50	.2963
Error	52	6	8.67		
Total	204	11			

Factor A: Using Excel, the p-value corresponding to $F = 1.38$ is .2846. Because p-value $> \alpha = .05$, Factor A (translator) is not significant.

Factor B: Using Excel, the p-value corresponding to $F = 6.57$ .0308. Because p-value $\leq \alpha = .05$, Factor B (language translated) is significant.

Interaction: Using Excel, the p-value corresponding to $F = 1.50$ is .2963. Because p-value $> \alpha = .05$, Interaction is not significant.

44.

		Factor B Manual	Factor B Automatic	Factor B Means
Factor A	Machine 1	$\bar{x}_{11} = 32$	$\bar{x}_{12} = 28$	$\bar{x}_{1.} = 30$
	Machine 2	$\bar{x}_{21} = 21$	$\bar{x}_{22} = 26$	$\bar{x}_{2.} = 23.5$
Factor B	Means	$\bar{x}_{.1} = 26.5$	$\bar{x}_{.2} = 27$	$\bar{\bar{x}} = 26.75$

Step 1

$$SST = \sum_i \sum_j \sum_k \left(x_{ijk} - \bar{\bar{x}}\right)^2 = (30 - 26.75)^2 + (34 - 26.75)^2 + \cdot\cdot\cdot + (28 - 26.75)^2 = 151.5$$

Step 2

$$SSA = br\sum_i \left(\bar{x}_{i\cdot} - \bar{\bar{x}}\right)^2 = 2\ (2)\ [\ (30 - 26.75)^2 + (23.5 - 26.75)^2\ ] = 84.5$$

Step 3

$$SSB = ar\sum_j \left(\bar{x}_{\cdot j} - \bar{\bar{x}}\right)^2 = 2\ (2)\ [\ (26.5 - 26.75)^2 + (27 - 26.75)^2\ ] = 0.5$$

Step 4

$$SSAB = r\sum_i \sum_j \left(\bar{x}_{ij} - \bar{x}_{i\cdot} - \bar{x}_{\cdot j} + \bar{\bar{x}}\right)^2 = 2[(32 - 30 - 26.5 + 26.75)^2 + \cdot\cdot\cdot + (26 - 23.5 - 27 + 26.75)^2] = 40.5$$

Step 5

SSE = SST – SSA – SSB – SSAB = 151.5 – 84.5 – 0.5 – 40.5 = 26

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F	p-value
Factor A	84.5	1	84.5	13	.0226
Factor B	.5	1	.5	.08	.7913
Interaction	40.5	1	40.5	6.23	.0671
Error	26	4	6.5		

Total	151.5	7

Factor A: Using Excel, the p-value corresponding to $F = 13$ is .0226. Because p-value $\leq \alpha = .05$, Factor A (machine) is significant.

Factor B: Using Excel, the p-value corresponding to $F = .08$ is .7913. Because p-value $> \alpha = .05$, Factor B (loading system) is not significant.

Interaction: Using Excel, the p-value corresponding to $F = 6.23$ is .0671. Because p-value $> \alpha = .05$, Interaction is not significant.