Intermediate Statistical Investigations 1st edition Tintle Solutions Manual

SKU: 9781119634522-SOLUTIONS

CHAPTER 2

Controlling Additional Sources of Variation

Section 2.1

2.1.1 **B.**

2.1.2 **D.**

2.1.3 **B.**

2.1.4 **B.**

2.1.5 **C.**

2.1.6 **A.**

2.1.7

a. See the following table.

Source	DF	Sum of Squares	Mean Squares	*F*
Method	1	30	30	15
Person	10	50	5	
Error	10	20	2	
Total	21	100		

b. 0.80.

c. 50%.

d. The *F*-statistic is the most helpful. As that number is much greater than 4, there is strong evidence of an association.

2.1.8

a. **i.** Because most of the lines connecting the points don't appear to be slanted in one particular direction, the pairing doesn't appear to help much.

ii. Because the standard deviation of the differences is the larger (1.861) than either of the two group standard deviations (1.522, 1.400), the pairing doesn't appear to help much.

b. Because there doesn't appear to be any linear pattern in the data and the R^2 value is so small, there is not much of an association between words memorized when exercising and not, so pairing did not help much.

2.1.9

a. **i.** Because most of the lines connecting the points appear to be slanted in a particular direction (upper right to lower left), the pairing does appear to help.

ii. Because the standard deviation of the differences (0.088) is much smaller than either of the two group standard deviations (0.26, 0.273), the pairing does appear to help.

b. As there does appear to be a linear pattern in the data and the R^2 value is very large, there is an association between the times for the narrow and wide-angle paths, so pairing did help.

2.1.10

a. 3.2.

b. $SSTotal = 61.2$.

c. $4 - 3.2 = 0.8$, $2.4 - 3.2 = -0.8$.

d. $10(0.8)^2 + 10(-0.8)^2 = 12.8$.

e. $SSperson = 40.2$.

f. $SSError = 61.2 - 12.8 - 40.2 = 8.2$.

g. See table below; $R^2 = (12.8 + 40.2)/61.2 = 0.867$, so 86.7%.

Source	DF	Sums of squares	Mean squares	*F*
Condition	1	12.8	12.8	14.1
Person	9	40.2	4.47	4.91
Error	9	8.2	0.91	
Total	19	61.2		

2.1.11

a. 8.

b. 74.

c. −2, 2.

d. 40.

e. 34.

f. The percentage of total variation in the scores that is left unexplained is 54.1%. See the following table.

Source	DF	Sum of Squares	Mean Squares	*F*
Environment	1	40	40	9.412
Error	8	34	4.25	
Total	9	74		

g. See the following table.

Person	1	2	3	4	5
Distracting Score	6	5	8	4	7
Quiet Score	10	8	14	8	10
Average Score	8	6.5	11	6	8.5
Person Effect	0	−1.5	3	−2	0.5

h. 4% of the variation is left unexplained. See the following table.

Source	DF	Sum of Squares	Mean Squares	*F*
Environment	1	40	40	53.33
Person	4	31	7.75	
Error	4	3	0.75	
Total	9	74		

2.1.12 $$\textit{predicted score} = 8 + \begin{cases} 0 \textit{ if person 1} \\ -1.5 \textit{ if person 2} \\ 3 \textit{ if person 3} \\ -2 \textit{ if person 4} \\ 0.5 \textit{ if person 5} \end{cases} + \begin{cases} -2 \textit{ if distracting environment} \\ 2 \textit{ if quiet environment} \end{cases}.$$

2.1.13

a. 4.0485.

b. 158.855.

c. 0.0161, −0.0161.

d. 0.01613.

e. $50.177 \times 2 = 100.355$.

f. 58.48.

g. See the following table.

Source	DF	Sum of Squares	Mean Squares	*F*
Condition	1	0.01613	0.01613	0.0083
Person	30	100.355	3.345	
Error	30	58.339	1.944	
Total	61	158.855		

h. 0.01%.

i. 63.17%.

j. No, not even close. The *F*-statistic is almost 0, nowhere near being larger than 4.

2.1.14

a. See Table 2.1.14a.

b. 16.5% explained by media and 83.5% unexplained.

c. See Table 2.1.14c.

d. 68.2% and 31.8%.

Source	DF	Sums of squares	Mean squares	*F*
Media	1	0.74	0.74	6.721
Error	34	3.74	0.11	
Total	35	4.48		

Table 2.1.14a

Source	DF	Sums of squares	Mean squares	*F*
Media	1	0.74	0.74	8.81
Person	17	2.31	0.136	1.62
Error	17	1.43	0.084	
Total	35	4.48		

Table 2.1.14c

2.1.15

a. H_0: $\mu_d = 0$, H_a: $\mu_d \neq 0$, where μ_d is the long-term mean difference in headway variability between using Facebook and Instagram.

b. 0.030.

c. Yes, there does seem to be a significant difference because 14 of the 18 differences are positive.

d. The p-value should be approximately 0.003.

e. $t = \dfrac{0.03 - 0}{0.037/\sqrt{18}} \approx 3.44$.

f. p-value = 0.0031.

g. We have strong evidence that the long-run mean difference in headway variability is not zero, with Facebook having the significantly larger variability.

2.1.16

a. (0.0828, 0.4907).

b. Yes, there is strong evidence that the mean time headway variability differs between drivers scrolling through Facebook and Instagram because the 95% confidence interval is completely positive.

2.1.17

a. H_0: $\mu_d = 0$, H_a: $\mu_d \neq 0$, where μ_d is the long-term mean difference in Stroop test times between listening to music and not listening to music.

b. 1.645.

c. Yes, there does seem to be a significant difference as 32 of the 38 of differences are positive.

d. The p-value should be approximately 0.0002.

e. $t = \dfrac{1.654 - 0}{2.657/\sqrt{38}} \approx 3.84$.

f. p-value = 0.0005.

g. We have strong evidence that the long-run mean difference in Stroop test time is not zero, with the no music environment giving significantly larger scores.

2.1.18

a. (0.7803, 2.5268).

b. Yes, there is strong evidence that there is a difference in mean Stroop test time because the 95% confidence interval is completely positive.

Section 2.2

2.2.1 B.

2.2.2 C.

2.2.3 D.

2.2.4 A.

2.2.5 D.

2.2.6 A.

2.2.7 D.

2.2.8

a. Sources of Variation Diagram:

Observed variation in: Racing Speed	Sources of explained variation	Sources of unexplained variation
Inclusion criteria • Racing dogs *Design* • Each dog received each diet in a random order	• Diets • Dog (breed, metabolism, sleep, race experience, etc.)	• Track conditions • Weather conditions • etc.

b. **i.** False.
ii. True.
iii. False.
iv. True.
v. True.

2.2.9 The null distribution of F-statistics created by shuffling all the responses (completely randomized) tends to have a larger standard deviation.

2.2.10

Source	DF	Sums of squares	Mean squares	F
Pan	1	4.5	4.5	22.5
Brand	1	2	2	10
Error	5	1	0.2	
Total	7	7.5		

2.2.11 The F-statistic is lower because the brand effect is smaller than the pan effect.

Source	DF	Sums of Squares	Mean squares	F
Brand	1	2	2	10
Pan	1	4.5	4.5	22.5
Error	5	1	0.2	
Total	7	7.5		

2.2.12

a. Randomly put 8 males in each group and 4 females in each group.

b. Randomly assign 3 of the 6 women over 40 to each group, and 3 of the 6 men over 40 to each group.

		Treatment A	Treatment B
Male	40 or older	3	3
	Under 40	5	5
Female	40 or older	3	3
	Under 40	1	1

2.2.13

a. See below.

Source	DF	Sums of squares	Mean squares	F
Detergent	1	0.28	0.28	0.085
Error	6	19.97	3.33	
Total	7	20.25		

b. 1.38%; F-statistic is 0.085.

c. 1.5625, −1.5625.

d. Polyester.

e. 19.531.

f. See below

Source	DF	Sums of squares	Mean squares	F
Detergent	1	0.28	0.28	3.182
Material	1	19.53	19.53	
Error	5	0.44	0.088	
Total	7	20.25		

g. 96.44%.

h. No, the percentage of variation in stain ratings explained by detergent type has not changed; Yes, the F-statistic has increased.

i. A large proportion of the total variation can be explained by material type which reduces the amount of unexplained variation (*SSError*). Because the denominator (*MSError*) is now so much smaller, the F-statistic is larger.

2.2.14
$$\textit{predicted stain rating} = 3.25 + \begin{bmatrix} -0.1875 \textit{ if Name Brand} \\ 0.1875 \textit{ if Store Brand} \end{bmatrix} + \begin{bmatrix} -1.5625 \textit{ if Polyester} \\ 1.5625 \textit{ if Cotton} \end{bmatrix}.$$

2.2.15

a. See the following table.

Source	DF	Sum of Squares	Mean Squares	F
Treatment	1	553.52	553.52	1.537
Error	10	3600.71	360.07	
Total	11	4154.23		

b. 13.3%; the *F*-statistic is 1.537.

c. −15.21, 15.21.

d. Non-organic.

e. 2775.521.

f. See the following table.

Source	DF	Sum of Squares	Mean Squares	*F*
Treatment	1	553.52	553.52	6.04
Type	1	2775.52	2775.52	
Error	9	825.39	91.71	
Total	11	4154.23		

g. 66.8%.

h. No, the percentage of variation in ripening time explained by wrapping has not changed; Yes, the *F*-statistic has increased.

i. A large proportion of the total variation can be explained by banana type and this reduces the amount of unexplained variation (*SSError*). Because the denominator (*MSError*) is now so much smaller, the *F*-statistic is larger.

2.2.16

$$\textit{predicted overripening time} = 48.54 + \begin{bmatrix} -6.79 \textit{ if not wrapped} \\ 6.79 \textit{ if wrapped} \end{bmatrix} + \begin{bmatrix} -15.21 \textit{ if organic} \\ 15.21 \textit{ if non-organic} \end{bmatrix}.$$

2.2.17

a. All the cards are being shuffled and 12 are randomly placed in each group. The graph displays a dot that is the proportion of males in group 1 minus the proportion of males in group 2.

b. The mean should be about 0 and the SD should be about 0.20. A mean of zero makes sense because, on average, you should have the same proportion of males in each group, so the difference in group proportions should be 0.

c. The mean should be about 0 (it will not be quite as close to 0 as the mean of the previous distribution) and the SD should be about 1.38.

d. Now the applet shuffles the 16 blue (male) cards and random puts 8 in each group, then randomly shuffles the 8 red (female) cards and randomly puts 4 in each group. This results in 8 blue and 4 red cards in each group after every shuffle.

e. The mean and SD are both 0. Because we are blocking on sex, we are forcing the proportion of males to be the same in each group so that all the differences in the proportions of males are 0.

f. The SD is about 0.85 and this is smaller than that from part (c) when we weren't blocking on sex.

g. Height is associated with sex because males, on average, are taller than females.

2.2.18

a. **i.** See below

Source	DF	Sums of squares	Mean squares	*F*	p-value
Model	5	749.50	149.90	0.038	0.9991
Error	18	71285.00	3960.28		
Total	23	72034.50			

ii. 749.5/72034.5 = 0.0104, about 1%.

iii. 71285/72034.5 = 0.9896, about 99%.

iv. The p-value is 0.9991 so there is not any evidence of a treatment effect.

b. **i.** See below

Source	DF	Sums of squares	Mean squares	*F*	p-value
Model	3	47852.83	15950.94	13.193	0.0001
Error	20	24181.67	1209.08		
Total	23	72034.50			

ii. 47852.83/72034.5 = 0.6643, about 66%.

iii. 24181.67/72034.5 = 0.3357, about 34%.

iv. The p-value is 0.0001 so there is very strong evidence of a block effect.

c. **i.** See Table 2.2.18c.i. This is the same ANOVA table as question 2.2.18(a)i.

ii. See Table 2.2.18c.ii.

iii. The sum of squares for treatment is the same as it was when ignoring the block, and the sum of squares for bock is the same as when ignoring the treatment. *SSModel* = 749.50 + 47852.83 = 48602.33 and 48602.33/72034.50 = 0.6747, about 67%.

iv. 23432.17/72034.50 = 0.3253, about 33%.

v. The blocking variable did explain a significant amount of variability in the apple weights, but not enough to result in a significant treatment effect.

vi. The apple orchard farmer should find out the conditions in block #4 where the apple production was highest and see whether those conditions can be replicated throughout the orchard. The type of ground cover did not have a significant effect on the apple production, so use whatever method was least expensive.

Source	DF	Sums of squares	Mean squares	*F*	p-value
Treatments	5	749.50	149.90	0.04	0.9991
Error	18	71285.00	3960.28		
Total	23	72034.50			

Table 2.2.18c.i

Source	DF	Sums of squares	Mean squares	*F*	p-value
Treatments	5	749.50	149.90	0.10	0.9915
Blocks	3	47852.83	15950.94	10.21	0.0007
Error	15	23432.17	1562.14		
Total	23	72034.50			

Table 2.2.18c.ii

2.2.19

a. Experimental units: the bananas; explanatory variable: type of treatment; response variable: hours until over-ripened.

b. See the following table.

Observed variation in: Ripening time	**Sources of explained variation**	**Sources of unexplained variation**
Inclusion criteria Bananas from same store *Design* Randomly assign treatment to each type of banana	• Treatment: bag, wrap, both, neither	• Type of banana • How long in store

c. 5511.21/8854.65 ≈ 0.62 so 62% of the variability in time until over-ripened is explained by the treatment and 3343.44/8854.65 ≈ 0.38 so 38% is still left unexplained.

d. $F = 4.4$.

e. p-value ≈ 0.06 (theory-based p-value = 0.0417).

f. The average hours until over-ripe for the 4 organic bananas is 53.3 hours. The average hours until over-ripe for all the bananas is 73.9 hours. The effect of the organic bananas is 53.3 − 73.9 = −20.6 hours.

g. **i.** The F-statistic has increased to 17.3. This makes sense because by adjusting the data by the effects for each banana type, the values are less variable which decreases the *SSError*, but the treatment means haven't changed, so the *SStreatment* will stay the same;

ii. p-value = 0.002 (theory-based p-value = 0.0023);

iii. Adjusting for the type of banana shows a stronger association between the treatment and the hours until over-ripened.

h. **i.** $F = 4.4$;

ii. The data are shuffled one block at a time;

iii. p-value ≈ 0.002;

iv. This is the same as the p-value for the adjusted data.

Section 2.3

2.3.1 B.

2.3.2 A.

2.3.3 A. and B.

2.3.4 D.

2.3.5 A.

2.3.6

a. False. If there is covariation then that will also be part of the total sum of squares.

b. False; If there is covariation then that will also be part of the sum of squares model.

c. False. If variable 2 is not associated with the response variable, adding it won't change the p-value.

d. False. If there is covariation between variables 1 and 2, then the sum of squares for variable 1 will change.

e. True.

2.3.7 *SSrace* is the sum of squares for race when education is also in the model.

2.3.8 A + B − C.

2.3.9 C − (A + B)

2.3.10

a. B.

b. D.

c. B + C + D.

d. C.

e. A.

2.3.11

a. No, age is not associated with Breed because each breed in the sample has 5 young and 5 mature cows.

b. There is not strong evidence of an association (p-value = 0.5276). *SSModel* = 0.21; 0.00414.

c. There is strong evidence of an association (p-value < 0.0001). *SSModel* = 34.32; 0.677.

d.

Source	DF	Adjusted SS	Adjusted MS	Adjusted F
Breed	4	34.321	8.580	49.907
Age	1	0.207	0.207	1.204
Error	94	16.161	0.172	
Total	99	50.689		

e. Yes, the adjusted SS match the *SSModel* numbers from parts (b) and (c).

f. Yes, the adjusted SS for breed, age, and error sum to *SSTotal* because we have the same proportion of mature or young cows in each breed and therefore are balanced.

g. *SScovariation* = 0

h. **i.** Pro: Stratified random samples can ensure adequate sample sizes in certain groups and can lead to minimized covariation which can improve statistical power. Con: Stratified samples, however, will no longer be representative of the population (overall), although particular strata in the sample may be representative of those same strata in the population (depending on the sampling strategy).

ii. There will be no covariance if data are gathered such that variable 1 is not associated with variable 2 (e.g., the proportions of the levels for variable 2 are the same for all levels of variable 1).

2.3.12

a. **i.** $R^2 = 0.677$.

ii. Ayrshire mean = 4.06%, Canadian mean = 4.44%, Guernsey mean = 4.95%, Holstein mean = 3.67%, Jersey mean = 5.29%.

b. $R^2 = 0.681$.

c. The R^2 values are rather similar because age is not very strongly associated with butterfat percentage.

d. The mean butterfat percentages are the same whether age is part of the model or not. This is because we have the same number of cows in each age–breed group. Everything is balanced just like it would be in an experiment.

2.3.13

a. There is not strong evidence of an association (p-value = 0.1532). *SSModel* = 198.35; 0.0244.

b. There is strong evidence of an association (p-value = 0.0000). *SSModel* = 2429.83; 0.299.

c.

Source	DF	Adj SS	Adj MS	Adj *F*
Sex	1	2389.35	2389.35	35.41
Allergy	1	157.88	157.88	2.340
Error	82	5533.56	67.48	
Total	84	8121.26		

d. No, the adjusted SS do not match the *SSModel* numbers found in parts (b) and (c).

e. No, because we do not have the same proportion of males and females with allergies, the adjusted *SSsex*, *SSallergy*, and *SSError* do not sum to *SSTotal*.

f. *SScovariation* = 8121.26 − 5533.56 − 157.88 − 2389.35 = 40.47.

g. No. 2.44% + 29.9% ≠ 31.36%.

2.3.14

a. 1 − (5533.56/8121.26) ≈ 0.3386 so 31.86%.

b. 2389.35/ 8121.26 ≈ 0.2942 so 29.42%.

c. 157.88/8121.26 ≈ 0.0194 so 1.94%.

d. 0.3186 − 0.2942 − 0.194 = 0.005, so 0.50%.

e. 5533.56/8121.26 ≈ 0.6814 so 68.14%.

f. Yes, 29.42% + 1.94% + 0.5% + 68.14% = 100%.

2.3.15

a. **i.** $R^2 = 0.0244$;

ii. no allergy mean height = 171.08 cm, allergy mean height = 167.94 cm.

b. **i.** $R^2 = 0.319$;

ii. no allergy = 168.86 − (−1.4) = 170.26 cm, allergy = 168.86 − 1.4 = 167.46 cm.

c. The R^2 is much larger when sex is added to the model because height is strongly associated with a person's sex.

d. The means from part (b) when sex is added to the model are lower than those from part (a). There were more males in the sample than females so when average height is adjusted for a person's sex these means decrease.

2.3.16

a. The average GPA of those who ate breakfast is 3.53 and it is 3.35 for those who didn't. With a p-value of 0.0350, there is strong evidence of an association between eating breakfast and GPA. Higher GPAs are seen with those who eat breakfast.

b. For these data, a person's sex is strongly associated with both GPA and eating breakfast. This means there is a fairly large covariation so when a person's sex is accounted for, we don't have strong evidence of an association between eating breakfast and GPA.

2.3.17

a. **i.** $R^2 = 0.05576$.

ii. Breakfast mean GPA = 3.53; No breakfast mean GPA = 3.350.

iii. Yes, with a p-value of 0.035 there is strong evidence of an association between eating breakfast and GPA.

b. **i.** $R^2 = 0.13535$.

ii. Breakfast mean GPA = 3.48; No breakfast mean GPA = 3.35.

iii. No, p-value is 0.13, so there is not a significant difference between the GPAs of those who ate breakfast and those who didn't.

iv. 70.0% of the females ate breakfast and 46.7% of the males ate breakfast.

v. The mean GPA for females is 3.55 and the mean GPA for males is 3.30.

c. The R^2 values are quite different. R^2 increased from about 0.055 to 0.135 by adding sex to the model. This means that sex is associated with eating breakfast which we can see in how different the percentages that ate breakfast are (70.0% and 46.7%).

d. The means got closer together when adjusted for a person's sex. The no breakfast mean increased because most of the males (with lower GPAs) were in that group so when it was adjusted for sex the mean increased. The breakfast mean decreased because most of the females (with higher GPAs) were in that group, so when it was adjusted for sex the mean decreased.

2.3.18

a. 1.42.

b. 1.42/10.46 × 100% = 13.58%.

c. 0.83/10.46 × 100% = 7.93%.

d. 0.28/10.46 × 100% = 2.68%.

e. (1.42 − 0.28 − 0.83)/10.46 × 100% = 2.96%.

f. 9.04/10.46 × 100% = 86.42%.

g. Close; they sum to 99.99% because of a little bit of rounding error.

2.3.19

a. Females have a higher mean GPA (3.55 vs. 3.30).

b. Those who ate breakfast have a higher mean GPA (3.53 vs. 3.35).

c. Yes, 35/50 or 70% of the females ate breakfast and only 14/30 or 46.7% of the males ate breakfast.

d. When the GPAs are adjusted for sex, the red dots move up and the blue dots move down. Because male GPAs were lower, on average, they increase and the female GPAs decrease when the sex effects are subtracted from each GPA.

e. The adjusted line is closer to horizontal. The slope of this (purple) line is smaller than the slope of the unadjusted line because the female GPAs were adjusted downward, and there are many more female dots in the breakfast column (more females ate breakfast than didn't) than male dots, so the column moves downward, on average.

2.3.20

a. Those who played a high school sport are taller in the sample than those who didn't (67.58 inches vs. 66.15 inches.) The p-value for testing this is 0.1393 so there is not strong evidence of an association.

b. **i.** When the heights are adjusted for sex, the blue dots (female heights) move up and the red dots (male heights) move down. The mean heights both went up because there are more females than males in the sample, but now both are close to the same (68.33 in vs. 68.06 in);

ii. The adjusted line is closer to horizontal. There are more females than males so both means increased, but a larger proportion of the no sport group is female so that mean increased more to level the line out.

c. The HS_Sport p-value is now 0.7047. By adjusting for sex, the mean heights were closer together, so the p-value increased from what it was when sex was not in the model.

2.3.21

a. **i.** The average alcohol consumption for those with a history of depression is 66.13 grams and those without 83.27 g, giving a difference of 17.14 g.

ii. *SSdepression* = 45,555; $R^2 = 0.0088$.

iii. Yes, there is strong evidence of an association because the p-value is 0.0106. Because this is an observational study, the small p-value is not evidence of a cause-and-effect relationship.

b. **i.** The average alcohol consumptions for males is 104.04 g and for females it is 46.85 g.

ii. *SSsex* = 599,712; $R^2 = 0.1160$.

iii. Yes, there is strong evidence of an association as the p-value is less than 0.0001.

c. **i.**

	Females	Males
History of depression	137 (40.9%)	84 (20.7%)
No history of depression	198 (59.1%)	321 (79.3%)
Total	335	405

ii. Yes, there seems to be an association because 40.9% of the females had a history of depression and only 20.7% of the males did.

iii. Yes, sex is a confounding variable because it is associated with both depression and alcohol consumption.

d. **i.**

	Mean Alcohol Consumption	
	Females	**Males**
History of depression	41.11	106.94
No history of depression	50.83	103.29

ii. The difference for females is 9.72 g and for males it is −3.65 g. The difference was 17.14 before so these differences are much smaller and for the males the direction of the association has even changed.

iii. *SSdepression* = 2,004. This is not the same as the SS from part (a), it is much smaller.

iv. Adjusted $R^2 = 0.0004$. It was 0.0088 in part (a), so it is smaller now.

v. There is not statistical significance because the p-value is 0.5697.

vi. Some of the variation in alcohol consumption that the model explained is explained by both depression and sex. This is the covariation. Most of the original variation represented by *SSdepression* has moved over to the variation represented by *SScovariation*. That is what made the *SSdepression* so small and the p-value so large.

2.3.22

a. **i.** The average number of cigarettes smoked for those who didn't finish high school is 6.86 and for those who did is 9.02 for a difference of 2.16 cigarettes;

ii. *SSeducation* = 136.78; 3.35%;

iii. There is strong evidence of an association because the p-value is 0.0316. Because this was an observational study, this is not evidence of a cause-and-effect relationship.

b. **i.** The average number of cigarettes for those less than 40 is 10.80 and for those 40 or more is 6.93;

ii. *SSage* = 482.26; 11.81%;

iii. There is strong evidence of an association because the p-value is about 0.

c. **i.**

	Less than 40 years old	40 or older
Finished High School	47 (92.16%)	49 (56.32%)
Didn't Finish HS	4 (7.84%)	38 (43.67%)
Total	51	87

ii. Because 92.2% of those less than 40 years old finished high school and only 56.3% of those 40 and older finished, there seems to be an association between education and age;

iii. Yes, age is a confounding variable because it is associated with both education and cigarettes.

d. **i.**

	Mean cigarette use	
	Less than 40 years old	**40 or older**
Finished High School	10.87	7.24
Didn't Finish HS	10.00	6.53

ii. For those less than 40 the difference is 10.87 – 10 = 0.87 and for those 40 or older it is 7.24 – 6.53 = 0.71. These are both much smaller than the original difference of 2.16;

iii. The *SSeducation* is 13.78 and it is much smaller than it was before adjusting for age;

iv. The percentage of variation explained by education is 0.34% which is much smaller than before adjusting for age;

v. There is not strong evidence of an association between Education and cigarette use now because the p-value is 0.47;

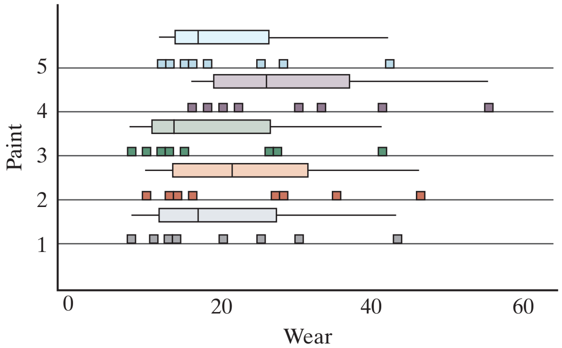
vi. Some of the variation cigarette use that the model explained is explained by both education and age. This is the covariation. Most of the original variation represented by *SSeducation* has moved over to the variation represented by *SScovariation*. That made the *SSeducation* so small and the p-value so large.

End of Chapter 2 Exercises

2.CE.1.

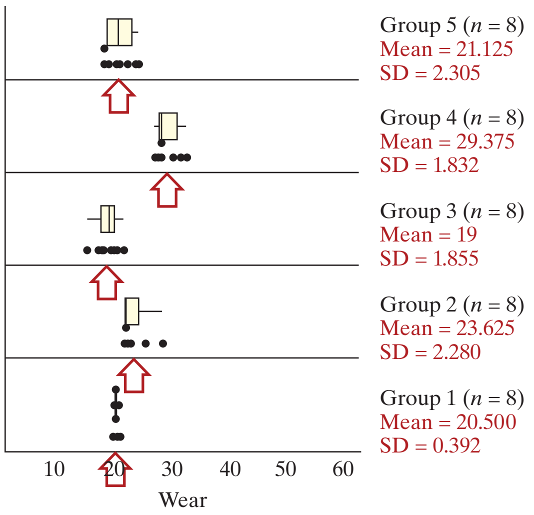
a. Five paints (explanatory variable) were randomly in eight different locations. Because random assignment to paints was carried out in each location, location is a blocking variable and hence this is a randomized block design. The response variable is the combined measure of wear.

b. The means differ (especially for paint 4) but there is still a lot of overlap in the distributions. The within group standard deviations are actually rather similar to the overall standard deviation, indicating we have not explained much of the variation in wear.



		Measure of Wear		
		n	**Mean**	**SD**
Treatment	**5**	8	21.13	10.15
	4	8	29.38	13.39
	3	8	19.00	11.34
	2	8	23.63	12.59
	1	8	20.50	11.72
Resid		40	0.00	11.89

c. The treatment means have not changed, but the standard deviations are **much** smaller (better to look at standard deviations than standard errors for this comparison). We see that paint type is statistically significant ($F = 30.42$, p-value < 0.0001) after adjusting for location.



Source	DF	Sum of Squares	Mean Squares	F	p-value
Paint	4	531.35	132.84	30.42	0.0000
Location	7	4826.37	689.48	157.92	0.0000
Error	28	122.25	4.37		
Total	39	5479.97			

d. We have random assignment within each location; the "location-adjusted" distributions still look reasonably well behaved (so we will assume normality in the treatment populations). However, Paint 1 does have a much smaller standard deviation than the others, perhaps indicating that after adjusting for location the variability in the paint types is not the same across the paint types.

e. Paint type 1 (the control) and Paint type 3 have significantly below average wear, and Paint type 4 has significantly above average wear. Clearly, Paint type 4 is best to maximize wear. We might not have detected this if we hadn't adjusted for location. And, of course, not adjusting for location would have been an incorrect analysis here.

2.CE.2

a. Observational units = 6 sections of a psychology class; Blocking variable = teacher; Factor = study method (3 levels).

b. Block effect for teacher A = 70 − 77.5 = −7.5 and block effect for teacher B = 80 − 77.5 = 7.5. Yes, we see that for all three study methods, teacher A has a lower average than teacher B. The teacher A sections score 7.5 points lower than the overall average, whereas the teacher B sections score 7.5 points higher than the overall average.

c. Yes, if we compute the three means for the Study Methods, we get 72.5, 82.5, 77.5. This would indicate the Small group + Individual group has the highest mean, and we see that it is higher than the other two groups for both teaching methods (75 > 70 > 65, and 90 > 85 > 80). Yes, if the scores are adjusted for the teacher effects, the study method means will stay the same. Adjusting for the block effects will reduce the variability of the scores within each study method, but it will not change the means. Because there will be less within study method variation in scores, the observed differences in means will actually be more statistically significant than before adjusting for the effects of the teachers.

d. See the following table.

Source	DF
Teacher	2 − 1 = 1
Study method	3 − 1 = 2
Error	2
Total	5

2.CE.3.

a. The experimental units are the platoons (or the commanders), because the treatment (type of commander) was applied to the platoon as a whole, not randomly assigned to individual soldiers. (Consequently, we computed the average to get one value per platoon.).

b. Explanatory Variable = Pygmalion or control; Blocking variable = company; so, you can think of this as two treatments with a blocking factor. There is a bit of replication because there are three platoons in most of the blocks, so there are two platoons getting the control treatment.

c. Null hypothesis: There is no difference in average score between these two treatments; Alternative hypothesis: There is a difference in the long-run average score. Test-statistic, $F = 3.52$, and p-value $= 0.0934$; After adjusting for the effect of company, there is moderate (but not strong) evidence (p-value = 0.0934) that there is a Pygmalion effect. (See the following ANOVA table.).

Source	DF	Sum of Squares	Mean Squares	*F*	p-value
Treatments	1	241.51	241.51	3.52	0.0934
Blocks	9	424.04	47.12	0.69	0.7080
Error	9	617.76	68.64		
Total	19	1283.32			

d. Using software, the 95% confidence interval for difference in mean score between the Pygmalion and control group (P – C) is: (1.80, 12.64). We are 95% confident that the mean score for the Pygmalion treatment will be 1.8 to 12.6 points higher than the mean score for the Control treatment, after controlling for company.

2.CE.4

a. No, the percentage of crews of size 4 is 50% for both crew leaders A & B. Similarly, the percentage of crews of size 6 is 50% for both crew leaders A & B.

b. $R^2 = 0.549$.

Source	DF	Sums of squares	Mean squares	*F*	p-value
CrewSize	1	231.13	231.13	7.30	0.0355
Error	6	189.88	31.65		
Total	7	421.00			

The effect of crew size 6 is +5.38 and the effect of crew size 4 is −5.38. This is statistically significant at the 5% level with p-value 0.0355.

c. $R^2 = 0.955$.

Source	DF	Adjusted SS	Adjusted MS	Adjusted *F*	p-value
Crew leader	1	171.13	171.13	45.63	0.0011
Crew size	1	231.12	231.12	61.63	0.0005
Error	5	18.75	3.75		
Total	7	421.0			

The effects (± 5.38) and sum of squares for crew size stay the same. The R^2 value for the two-variable ANOVA is much higher than for the one-variable ANOVA. This is because crew leader explained additional variability in productivity. Even though the degrees of freedom for error decreased when the blocking variable, crew leader was added, the *SSError* decreased by such a large amount that crew size adjusted for crew leader is now even more of a significant predictor of productivity (p-value = 0.0005).

Chapter 2 Investigation

1. This is an observational study, because the researchers did not manipulate or control any of the explanatory variables.

2. Each individual who was recruited in 1971.

3. Response variable: Total cholesterol – quantitative.

Explanatory variable: Whether or not take cholesterol medication regularly – categorical.

4.

a. Adult children of the 1948 cohort, recruited in 1971, and attended the 8th examination cycle in 2005–2008.

b. There are no sources of variation accounted for by the study design.

c. Age, family history, what other medication they take regularly, other medical conditions, etc.

d. See the following table.

Observed Variation in: Total cholesterol level	**Sources of explained variation**	**Sources of unexplained variation**
Inclusion criteria • Adult children of the 1948 cohort, • recruited in 1971, and • attended the 8th examination cycle in 2005–2008.	• Whether or not take cholesterol medication regularly	• Age • Family history • What other medication they take regularly • Other medical conditions • Unknown

5. Overall mean = 185.40, Overall SD = 40.62, *SSTotal* = 477,7736.7 *predicted total cholesterol level* = 185.40, *SE of residuals* = 40.62.

6. *predicted total cholesterol level*

$$= \begin{cases} 166.96 \text{ if cholesterol meds} \\ 199.29 \text{ if not} \end{cases}, \ \textit{SE of residuals} = 37.33.$$

Because the SE of residuals is reduced from 40.62 to 37.33 the predictions are more accurate if knowledge of whether or not someone is taking cholesterol medications regularly is used.

7. Least Squares mean estimate of cholesterol level = 183.12, Effect of taking cholesterol medication regularly = –16.17, Effect of not taking cholesterol medication regularly = 16.17

predicted total cholesterol level

$$= 183.12 + \begin{cases} -16.17 \text{ if cholesterol meds} \\ 16.17 \text{ if not} \end{cases}, \ \textit{SE of residuals} = 37.33.$$

8. *SSModel* = 742,431.0; *SSError* = 4,035,305.7; 0.155.

9. The observed difference in cholesterol levels between taking meds and not taking meds regularly is 32.34 mg/dL. This seems to be borderline significant with regard to practical significance.

10. Null hypothesis: There is no association between taking cholesterol meds regularly and cholesterol level; $\mu_{meds} = \mu_{no\ meds}$.

Alternative hypothesis: There is an association between taking cholesterol meds regularly and cholesterol level; $\mu_{meds} \neq \mu_{no\ meds}$, where μ_{meds} represents the mean cholesterol level in the population of those who take cholesterol meds regularly and $\mu_{no\ meds}$ represents the mean cholesterol level in the population of those who don't take cholesterol meds regularly.

11.

Source of Variation	DF	Sum of Squares	Mean Squares	*F* Ratio
Model	1	742431.0	742431	532.6
Error	2895	4035305.7	1394	**Prob > F**
Total	2896	4777736.7		**< .0001***

The validity conditions are met because the sample sizes in each group (1245 and 1646) are large, and the sample standard deviations for each group are similar (neither is more than twice as large as the other).

12.

Cholesterol Medication	Difference	SE Diff	Confidence interval
No − Yes	32.34	1.40	(29.59, 35.08)

We are 95% confident that the mean total cholesterol level is between 29.59 to 35.09 mg/dL higher among those who do not take cholesterol meds regularly compared to those who do take cholesterol meds regularly.

13. Yes, we have found evidence that total cholesterol level tends to be lower, on average, for those who regularly take cholesterol medication compared to those who don't because the ANOVA resulted in a large *F*-statistic (532.63) and a small p-value (< 0.0001), and the 95% confidence interval did not contain 0.

14. No, we cannot conclude causation from an observational study. Possible confounding variables include use of other medications, family history, etc.

15. Can either make a mosaic plot (such as the one shown for Solution 15) or a two-way table. There does appear to be an association between aspirin use and the use of cholesterol medication.

		Cholesterol Medicine		Total
		Yes	No	
Aspirin	Yes	778	512	1290
	No	468	1141	1609
Total		1246	1653	2899

16.

Analysis of Variance					
Source	DF	Sum of Squares	Mean Squares	*F*	p-value
Take_aspirin	1	253946.4	253946	162.5	< .0001
Error	2895	4523790.3	1563		
Total	2896	4777736.7			

The validity conditions are met because the sample sizes in each group (1290 and 1607) are large, and the sample standard deviations (39.07 and 39.90) for each group are similar (neither is more than twice as large as the other).

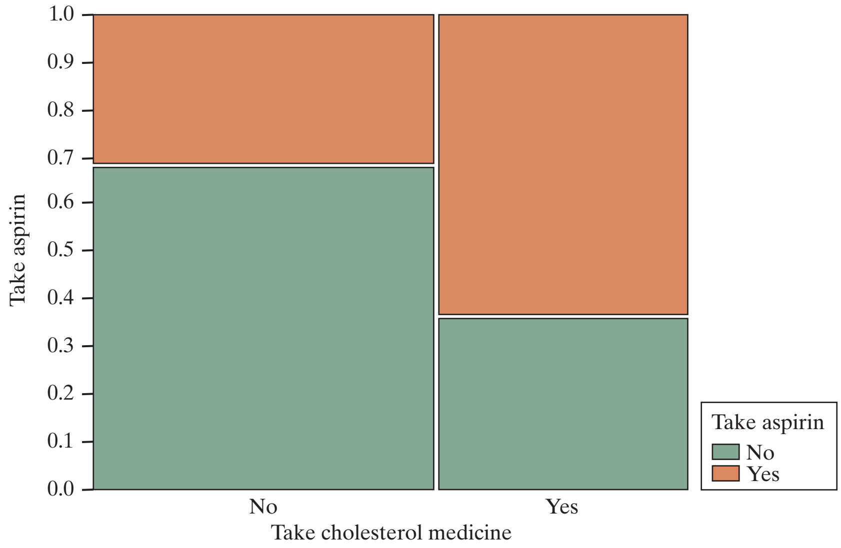
17. Yes, because the proportion of aspirin takers is higher among those who take cholesterol medication regularly, compared to those who do not take cholesterol meds regularly, aspirin use is a confounding variable. And, taking aspirin has a statistically significant association with total cholesterol level (p-value < 0.001).

18.

Source	DF	Sum of Squares	Mean Squares	*F*	p-value
Cholesterol medication use	1	549013.8	549013.8	399.7	< 0.0001
Aspirin use	1	60529.2	60529.2	44.1	< 0.0001
Error	2883	3974776.5	1373		
Total	2886	4777736.7			

The validity conditions are met because the sample sizes are very large and the standard deviations of each group are close together.

19. The *SScholesterol* is now smaller than before. The effect of taking cholesterol medication is –14.6—closer to zero—so a smaller effect in absolute value. Both the R^2 (0.115) and the *F*-statistic (399.7) are now



Solution 15

smaller. The *SE of residuals* = 37.06. The changes make sense because after adjusting for the effects of taking aspirin (which we saw has a statistically significant association with lower cholesterol levels), the effect of taking cholesterol medication on total cholesterol level are smaller. But, because use of aspirin and use of cholesterol medication jointly explain more variation that just cholesterol medication usage itself, the SE of the residuals is smaller.

20. *predicted total chol.* = 182.81 +

$$\begin{cases} -14.64 \textit{ if chol. med} \\ 14.64 \textit{ if not} \end{cases} + \begin{cases} -4.84 \textit{ if aspirin} \\ 4.84 \textit{ if not} \end{cases}, \textit{SE of residuals} = 37.06.$$

21.

Cholesterol Medication	Difference	Std Err Diff	Confidence interval
No – Yes	29.29	1.46	(26.41, 32.16)

We are 95% confident that the mean aspirin-use adjusted total cholesterol level is between 26.41 to 32.16 mg/dL higher among those who do not take cholesterol meds regularly compared to those who do take cholesterol meds regularly.

22. We have strong evidence (p-value < 0.0001) that after adjusting for use of aspirin those who take cholesterol medication regularly tend to have lower cholesterol levels, on average, compared to those who do not take cholesterol medication regularly; the mean aspirin-use adjusted total cholesterol level is between 26.41 to 32.16 mg/dL lower among cholesterol medication takers.

23.

a. The *SScholesterol* is now even smaller than before. Both the R^2 (0.106) and the *F*-statistic (397) are now slightly smaller. The 95% CI is narrower than before, involving smaller effects. The changes make sense because after adjusting for the effects of taking aspirin (which we saw has a statistically significant association with lower cholesterol levels) and sex, the effect of taking cholesterol meds on total cholesterol level are smaller.

b. The effects are slightly smaller which is consistent with the narrower confidence interval.

c. 182.19 − 14.13 − 3.37 − 9.99 = 154.7 mg/dL.

d. We are 95% confident that on average, females who take both aspirin and cholesterol medication will have a total cholesterol level between 152.19 and 157.22 mg/dL.

e. We are 95% confident that a female who takes both aspirin and cholesterol medication will have a total cholesterol level between 84.54 and 224.86 mg/dL.

f. A 95% confidence interval for the population mean response is narrower than a 95% prediction interval for an individual's response because the confidence interval has a smaller margin of error because the standard error of the sample mean is smaller than that of a single predicted response.

g. Based on the three-variable ANOVA, we should adjust for the effects of aspirin use and whether person is a male or female (p-values < 0.0001), because these both turned out to have a statistically significant association with total cholesterol level. Thus, we have strong evidence (p-value < 0.0001) that after adjusting for use of aspirin and whether the person is a male or a female, those who take cholesterol meds regularly tend to have lower cholesterol levels, on average, compared to those who do not take cholesterol meds regularly, with the mean adjusted total cholesterol level being between 25.47 and 31.04 mg/dL lower among cholesterol med takers. We still cannot conclude causation because there are other possible confounding variables such as age, family history, general health, etc. for which we have not yet made any adjustments.

24. Answers may vary; but possible follow-up questions to pursue include what other variables can be adjusted for to investigate the relationship between use of cholesterol medication and cholesterol level. We can also propose/conduct a randomized experiment that will help us isolate the effect of the cholesterol medication usage.