11.29 (a)

(a) $df\,A = r - 1 = 3 - 1 = 2$

$df\,B = c - 1 = 3 - 1 = 2$

(b) $df\,AB = (r-1)(c-1) = (3-1)(3-1) = 4$

(c) $df\,E = rc(n' - 1) = 3 \text{ x } 3 \text{ x } (4-1) = 27$

d) $df\,T = n - 1 = 35$

11.30

11.30

Source	*df*	*SS*	*MS*	*F*
Factor *A*	2	120	(b) 120÷ 2 = 60	60÷ 10 = 6
Factor *B*	2	110	(b) 110÷ 2 = 55	55÷ 10 = 5.5
Interaction, *AB*	4	(a) 540–120–110–270=40	(c) 40 ÷ 4 = 10	10÷ 10 = 1
Error *E*	27	270	270÷27=10	
Total *T*	35	540		

11.31

Source	*df*	*SS*	*MS*	*F*
Factor *A*	2	120	120÷ 2 = 60	(b) 60÷ 10 = 6
Factor *B*	2	110	110÷ 2 = 55	(c) 55÷ 10 = 5.5
Interaction, *AB*	4	540 – 120 – 110 – 270 = 40	40÷ 4 = 10	(a) 10÷ 10 = 1
Error, *E*	27	270	270÷ 27 = 10	
Total, *T*	35	540		

11.32 $F_{(2,\,27)} = 3.35$ $F_{(4,\,27)} = 2.73$

(a) Decision: Since $F_{calc} = 6.00$ is greater than the critical bound of 3.35, reject H_0. There is evidence of a difference among factor A means.

(b) Decision: Since $F_{calc} = 5.50$ is greater than the critical bound of 3.35, reject H_0. There is evidence of a difference among factor B means.

(c) Decision: Since $F_{calc} = 1.00$ is less than the critical bound of 2.73, do not reject H_0. There is insufficient evidence to conclude there is an interaction effect.

11.33 (a)

Source	*Df*	*SS*	*MS*	*F*
Factor *A*	2–1 = 1	18	18÷ 1 = 18	18÷ 2 = 9
Factor *B*	5–1 = 4	64	64÷ 4 = 16	16÷ 2 = 8
Interaction, *AB*	1•4 = 4	150 – 18 – 64 – 60 = 8	8÷ 4 = 2	2÷ 2 = 1
Error, *E*	2•5•3= 30	60	60÷ 30 = 2	
Total, *T*	39	150		

$F_{(1, 30)} = 4.1709$ $F_{(4, 30)} = 2.6896$

(b)(i) Decision: Since $F_{calc} = 9$ is greater than the critical bound of 4.1709, reject H_0. There is evidence of a difference among factor A means.

(ii) Decision: Since $F_{calc} = 8$ is greater than the critical bound of 2.6896, reject H_0. There is evidence of a difference among factor B means.

(iii) Decision: Since $F_{calc} = 1$ is less than the critical bound of 2.6896, do not reject H_0. There is insufficient evidence to conclude there is an interaction effect.

11.34

Source	***Df***	***SS***	***MS***	***F***
Factor *A*	2	2 x 80 = 160	80	80÷5 = 16
Factor *B*	8÷2 = 4	220	220÷4 = 55	11
Interaction, *AB*	8	8 x 10 = 80	10	10÷5 = 2
Error, *E*	30	30 x 5 = 150	55÷11 = 5	
Total, *T*	44	160+220+80+150 = 610		

11.35 $F_{(2, 30)} = 3.32$ $F_{(4, 30)} = 2.69$ $F_{(8, 30)} = 2.27$

(a) Decision: Since $F_{calc} = 16$ is greater than the critical bound of 3.32, reject H_0. There is evidence of a difference among factor A means.

(b) Decision: Since $F_{calc} = 11$ is greater than the critical bound of 2.69, reject H_0. There is evidence of a difference among factor B means.

(c) Decision: Since $F_{calc} = 2$ is less than the critical bound of 2.27, do not reject H_0. There is insufficient evidence to conclude there is an interaction effect.

11.36

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample	52.563	1	52.563	23.579	0.000	4.747
Columns	1.563	1	1.563	0.701	0.419	4.747
Interaction	3.063	1	3.063	1.374	0.264	4.747
Within	26.750	12	2.229			
Total	83.938	15				

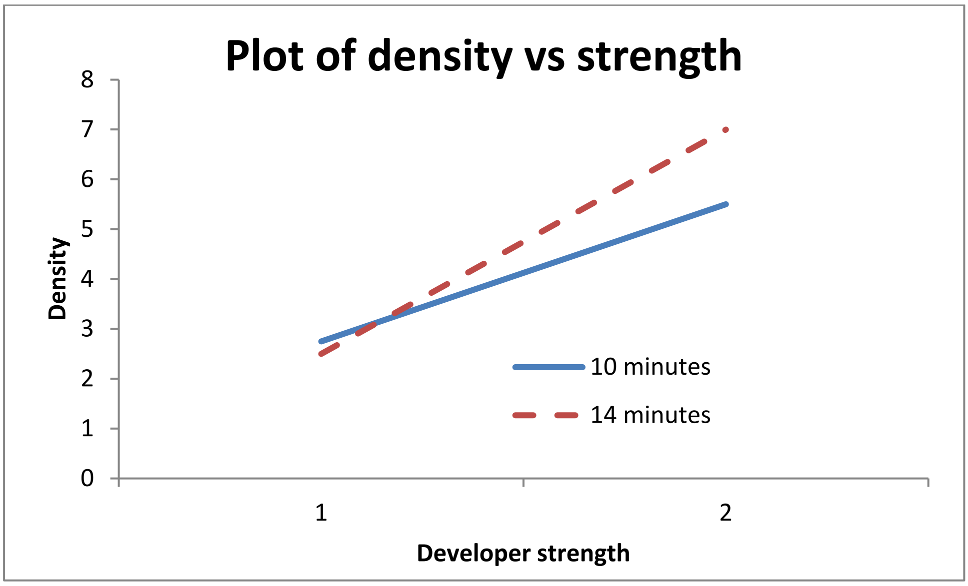
Excel Two-way ANOVA output:

(a) H_0: There is no interaction between development time and developer strength.
H_1: There is an interaction between development time and developer strength.
Decision rule: If $F_{calc} > 4.747$, reject H_0.
Test statistic: $F_{calc} = 1.374$.
Decision: Since $F_{calc} = 1.374$ is less than the critical bound of 4.747, do not reject H_0. There is insufficient evidence to conclude that there is any interaction between development time and developer strength.

(b) H_0: $\mu_1 = \mu_2$ H_1: $\mu_1 \neq \mu_2$
Decision rule: If $F_{calc} > 4.747$, reject H_0.
Test statistic: $F_{calc} = 23.58$.
Decision: Since $F_{calc} = 23.579$ is greater than the critical bound of 4.747, reject H_0. There is sufficient evidence to conclude that developer strength affects the density of the photographic plate film.

(c) H_0: $\mu_{10} = \mu_{14}$ H_1: $\mu_{10} \neq \mu_{14}$
Decision rule: If $F_{calc} > 4.747$, reject H_0.
Test statistic: $F_{calc} = 0.701$.
Decision: Since $F_{calc} = 0.701$ is less than the critical bound of 4.747, do not reject H_0. There is inadequate evidence to conclude that development time affects the density of the photographic plate film.

(d)



(e) At 5% level of significance, developer strength has a positive effect on the density of the photographic plate film while the developer time does not have any impact on the density. There is no significant interaction between developer time and developer strength on the density.

11.37

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample	528.125	1	528.125	24.143	0.008	7.709
Columns	5253.125	1	5253.125	240.143	0.000	7.709
Interaction	28.125	1	28.125	1.286	0.320	7.709
Within	87.500	4	21.875			
Total	5896.875	7				

(a) H_0: There is no interaction between the type of pasta and cooking time.
H_1: There is an interaction between the type of pasta and cooking time.
Decision rule: If $F_{calc} > 7.709$, reject H_0.
Test statistic: $F_{calc} = 1.286$
Decision: Since $F_{calc} = 1.286$ is less than the critical bound of 7.709, do not reject H_0. There is insufficient evidence to conclude that there is any interaction between the type of pasta and cooking time.

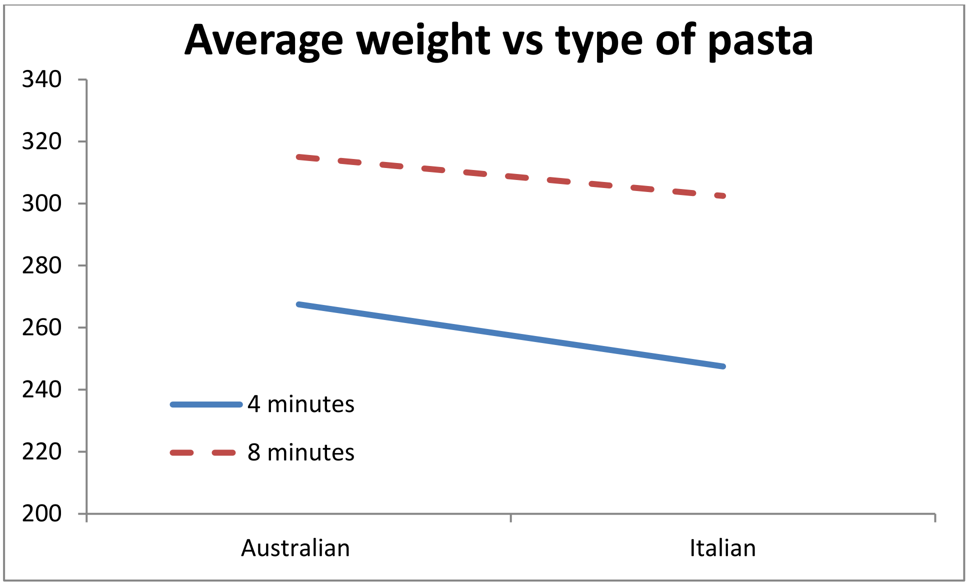
(b) H_0: $\mu_{\text{Australian}} = \mu_{\text{Italian}}$
H_1: $\mu_{\text{Australian}} \neq \mu_{\text{Italian}}$
Decision rule: If $F_{calc} > 7.709$, reject H_0.
Test statistic: $F_{calc} = 24.143$

Decision: Since F_{calc} = 24.143 is greater than the critical bound of 7.709, reject H_0. There is sufficient evidence to conclude that the type of pasta affects the weight of the pasta.

(c) H_0: $\mu_4 = \mu_8$ H_1: $\mu_4 \neq \mu_8$

Decision rule: If F_{calc} > 7.709, reject H_0.
Test statistic: F_{calc} = 240.143
Decision: Since F_{calc} = 240.143 is greater than the critical bound of 7.709, reject H_0. There is adequate evidence to conclude that cooking time affects the weight of the pasta.

(d)



(e) At 5% level of significance, cooking time has a positive effect on the weight of the pasta while the type of pasta also affects the weight of the pasta. There is no significant interaction between cooking time and the type of pasta on the weight of the pasta.

11.38

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample	1998.375	1	1998.375	15.361	0.001	4.414
Columns	6825.083	2	3412.542	26.231	0.000	3.555
Interaction	270.750	2	135.375	1.041	0.374	3.555
Within	2341.750	18	130.097			
Total	11435.958	23				

(a) H_0: There is no interaction between marketing technique and location of the market.
H_1: There is an interaction between marketing technique and location of the market.
$H_{0.05,2,18} = 3.555$
Since F_{calc} = 1.041 <3.555, do not reject H_0. There is no significant evidence of an interaction between marketing technique and location of the market.

(b) H_0: $\mu_A = \mu_B = \mu_C$
H_1: the means are different
Decision rule: If F_{calc} > 3.555, reject H_0.

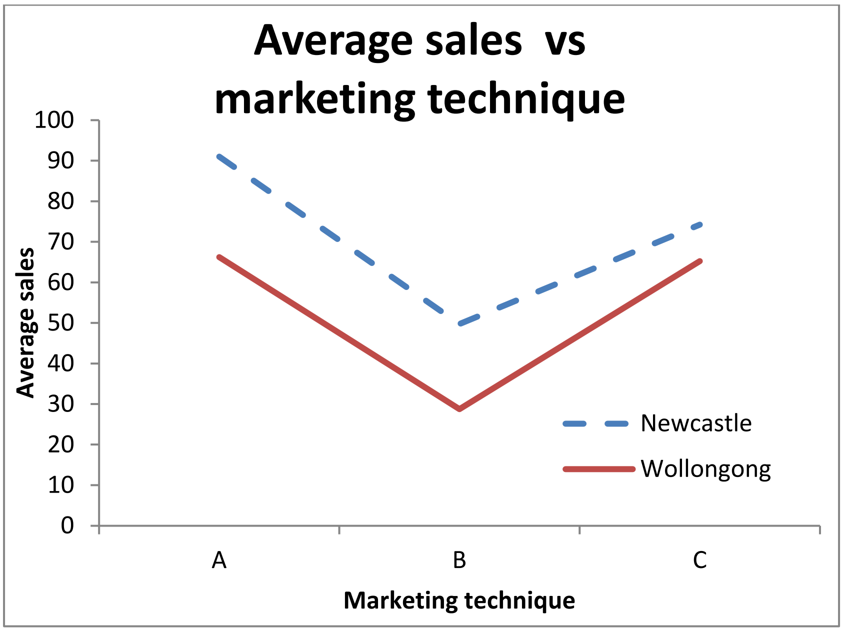
Test statistic: F_{calc} = 26.231
Decision: Since F_{calc} = 26.231 is greater than the critical bound of 3.555, reject H_0. There is adequate evidence to conclude that marketing technique affects the sales of a new product.

(c) H_0: $\mu_{\text{Newcastle}} = \mu_{\text{Wollongong}}$

H_1: $\mu_{\text{Newcastle}} \neq \mu_{\text{Wollongong}}$

Decision rule: If $F_{calc} > 4.414$, reject H_0.
Test statistic: F_{calc} = 15.361
Decision: Since F_{calc} = 15.361 is greater than the critical bound of 4.414, reject H_0. There is sufficient evidence to conclude that the location of the market affects the sales of a new product.

(d)



(e) The difference in the average sales of a new product depends both on the location of the market and on the marketing technique. The plot of the average sales versus marketing technique illustrates that on average sales are higher in Newcastle compared with Wollongong. The marketing technique A is the most successful. The marketing technique C provides the smallest difference between average sales in Newcastle and Wollongong.

11.39 The between-groups variance *MSB* represents variation among the means of the different groups. The within-groups variance *MSW* measures variation within each group.

11.40 In a completely randomised design, individual items in different samples are randomly and independently drawn. In a randomised block design, individual items in different samples are matched using common characteristics, or repeated measurements are taken to reduce within-group variation.

11.41 The completely randomised design evaluates one factor of interest, in which sample observations are randomly and independently drawn. The randomised block design also evaluates one factor of interest, but sample observations are divided into blocks according to common characteristics to reduce within-group variation. The two-factor factorial design evaluates two factors of interest and the interaction between these two factors.

11.42 The major assumptions of ANOVA are randomness and independence, normality and homogeneity of variance.

11.43 If the populations are approximately normally distributed and the variances of the groups are approximately equal, you should select the one-way ANOVA *F* test to examine possible differences among the means of *c* independent populations.

11.44 When the ANOVA has indicated that at least one of the groups has a different population mean than the others, you should use multiple comparison procedures for evaluating pairwise combinations of the group means. In such cases, the Tukey–Kramer procedure should be used to compare all pairs of means.

11.45 In the randomised block design, individual items in different samples are matched using common characteristics, or repeated measurements are taken to reduce within-group variation. In the two-factor factorial design, more than one observation can be obtained for each treatment combination to measure the interaction of two factors.

11.46 The ANOVA tests for differences across two or more population means whereas the Levene test is for the differences in population variances.

11.47 You should use the two-way ANOVA *F* test to examine possible differences among the means of each factor in a factorial design when there are two factors of interest that are to be studied and more than one observation can be obtained for each treatment combination (to measure the interaction of the two factors).

11.48 Interaction measures the difference in the effect of one variable for the different levels of the second factor. If there is no interaction, any difference between levels of one factor will be the same at each level of the second factor.

11.49 You can obtain the interaction effect and carry out an *F* test for its significance. In addition, you can develop a plot of the response for each level of one factor at each level of a second factor.

11.50

ANOVA: Two-Factor With Replication

SUMMARY	Sydney	Singapore	Wellington	London	Total
pre-crisis					
Count	5	5	5	5	20
Sum	37	43	33	32	145
Average	7.4	8.6	6.6	6.4	7.25
Variance	2.3	1.3	1.3	4.3	2.7236842
post-crisis					
Count	5	5	5	5	20
Sum	32	43	26	17	118
Average	6.4	8.6	5.2	3.4	5.9
Variance	1.3	0.3	0.7	1.3	4.5157894
Total					
Count	10	10	10	10	
Sum	69	86	59	49	
Average	6.9	8.6	5.9	4.9	

Variance	1.877777778	0.711111111	1.433333333	4.988888889

ANOVA

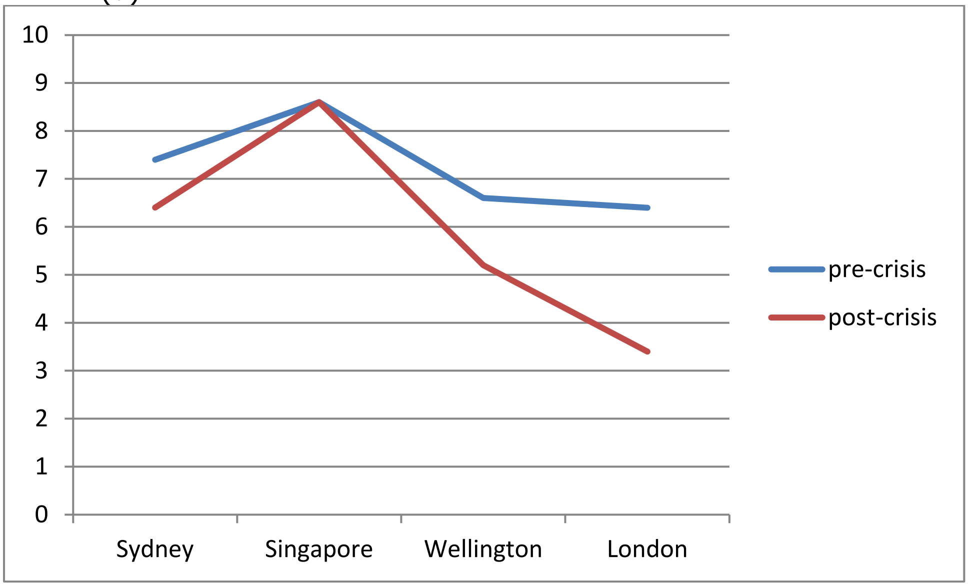
Source of Variation	SS	df	MS	F	P-value
Sample	18.225	1	18.225	11.390625	0.0019489
Columns	74.675	3	24.89166667	15.55729167	2.03661E-
Interaction	11.675	3	3.891666667	2.432291667	0.0830940
Within	51.2	32	1.6		
Total	155.775	39			

(a) H_0: There is no interaction between location and the GFC.
H_1: There is an interaction between location and the GFC.
Decision rule: If $F_{calc} > 2.901$, reject H_0.
Since $F_{calc} = 2.43 < 2.901$ do not reject H_0. There is not enough evidence to conclude that there is an interaction between location and GFC on consumer confidence.

(b) $H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4$
H_1: at least one mean differs
Decision rule: If $F_{calc} > 2.901$, reject H_0.
Since $F_{calc} = 15.557 > 2.901$, reject H_0. There is sufficient evidence to conclude mean consumer confidence differs by location.

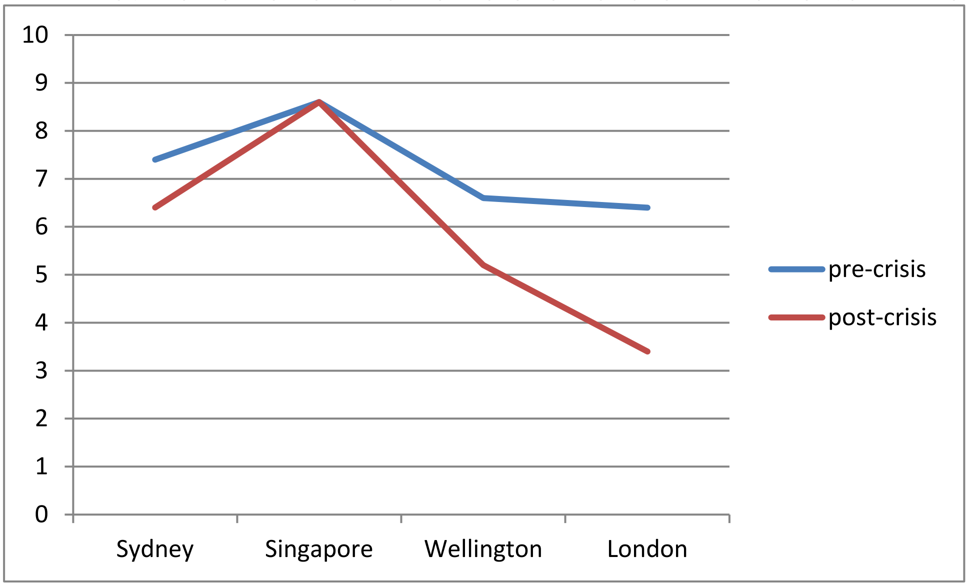
(c) $H_0 : \mu_1 = \mu_2$
H_1: at least one mean differs
Decision rule: If $F_{calc} > 4.149$, reject H_0.
Since $F_{calc} = 11.39 > 4.149$, reject H_0. There is adequate evidence to conclude that mean consumer confidence differs pre vs post_GFC.

(d)



(e) The Tukey procedure for pairwise comparisons is not used. The significant interaction makes the study of the main effects difficult.

(f) Both location and the GFC had individual (main) effects on mean consumer confidence separately. However, as there was no significant interaction effect there was not a combination of location and affected consumer confidence.



11.51 (a) H_0: $\sigma_1^2 = \sigma_2^2 = \sigma_3^2$

H_1: Not all σ_j^2 are the same.

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	0.07	2	0.035	0.07468	0.928383	3.682317
Within Groups	7.03	15	0.468667			
Total	7.1	17				

Since the p-value = 0.928 > 0.05, do not reject H_0. There is not enough evidence of a significant difference in the variances of the breaking strengths for the three air-jet pressures.

(b) H_0: $\mu_1 = \mu_2 = \mu_3$

H_1: At least one of the means differs.

Decision rule: If $F_{calc} > 3.89$, reject H_0.

ANOVA: Single Factor

ANOVA

Source of Variation	SS	df	MS	F	P-value	F crit
Between Groups	8.074	2	4.037	4.09	0.038	3.8853
Within Groups	14.815	15	0.988			
Total	22.889	17				

Test statistic: F_{calc} = 4.09

Decision: Since F_{calc} = 5.89 is greater than the critical bound of 3.68, reject H_0.

There is enough evidence to conclude that the mean breaking strengths differ for the three air-jet pressures.

(c)

Breaking strength scores under 30 psi are significantly higher than those under 50 psi.

(d) Other things being equal, use 30 psi.

11.52 (a)

Source of Variation	SS	Df	MS	F	P-value	F crit
Sample	112.5603	1	112.5603	30.4434	3.07E-06	4.113165
Columns	46.01025	1	46.01025	12.4441	0.001165	4.113165
Interaction	0.70225	1	0.70225	0.1899	0.665575	4.113165
Within	133.105	36	3.697361			
Total	292.3778	39				

(a) (a) H_0: There is no interaction between type of breakfast and desired time. H_1: There is an interaction between type of breakfast and desired time. Decision rule: If $F_{STAT} > 4.1132$, reject H_0.

Test statistic: $F_{STAT} = 0.1899$.
Decision: Since $F_{STAT} = 0.1899$ is less than the critical bound of 4.1132, do not reject H_0. There is insufficient evidence to conclude that there is any interaction between type of breakfast and desired time.

(b)

H_0: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 \neq 0$ Population 1 = Continental, 2 = American

Decision rule: If $F_{calc} >$ 4.1132, reject H_0. Test statistic: $F_{calc} = 30.4434$.
Decision: Since $F_{calc} = 30.4434$ is greater than the critical bound of 4.1132, reject H_0. There is sufficient evidence to conclude that there is an effect that is due to type of breakfast.

.

©H_0: $\mu_1 - \mu_2 = 0$ H_1: $\mu_1 - \mu_2 \neq 0$ Population 1 = Early, 2 = Late

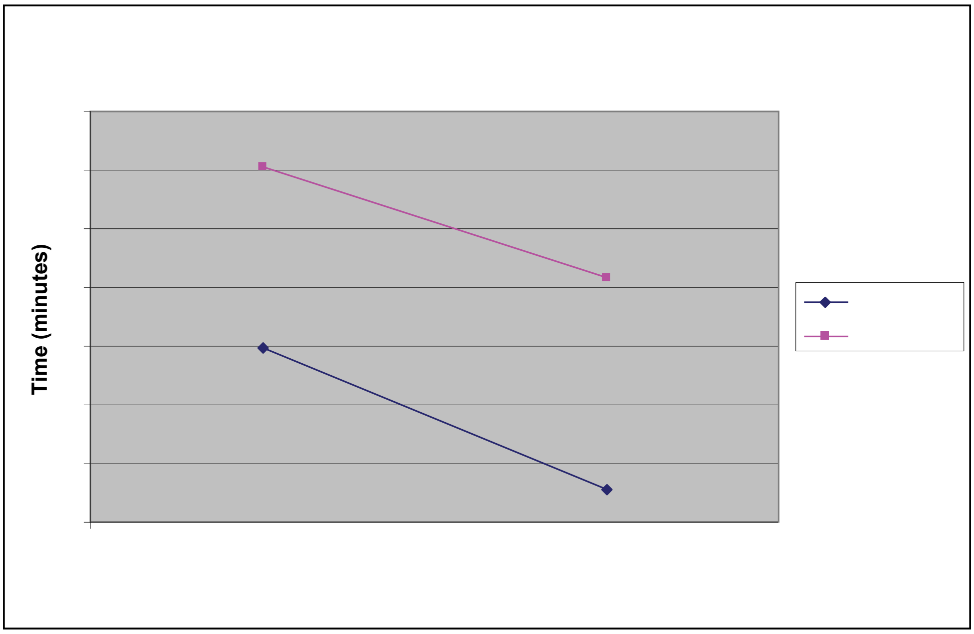
Decision rule: If $F_{calc} >$ 4.1132, reject H_0. Test statistic: $F_{calc} = 12.4441$.

Decision: Since $F_{calc} = 12.4441$ is greater than the critical bound of 4.1132, reject H_0. There is sufficient evidence to conclude that there is an effect that is due to desired time

(d)

c

.



.

(e) At the 5% level of significance, both the type of breakfast ordered and the desired time have an effect on delivery time difference. There is no interaction between the type of breakfast ordered and the desired time.

11.53 (a) H_0: There is no interaction between department and store.
H_1: There is an interaction between department and store.
Decision rule: If $F_{calc} > 3.885$, reject H_0.

ANOVA

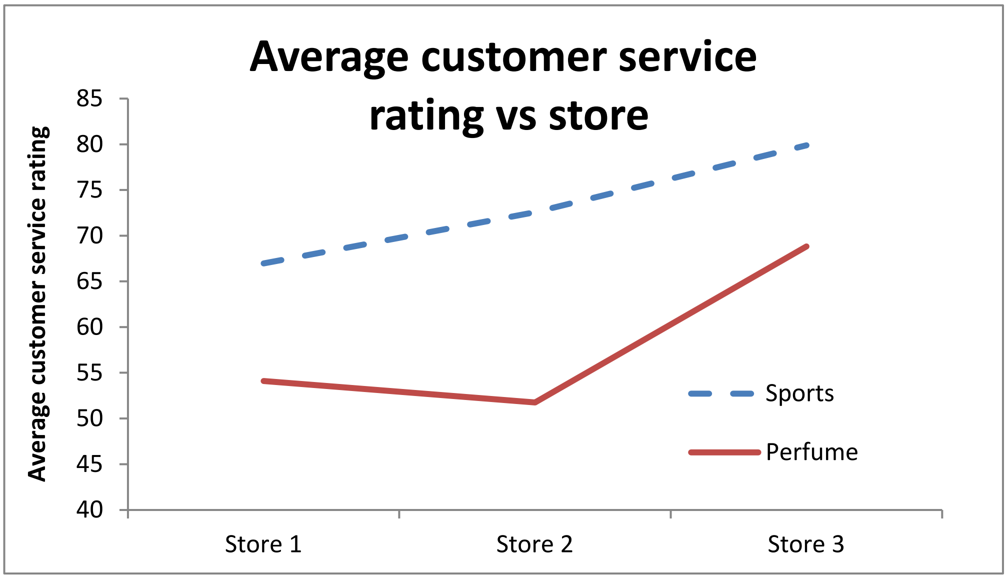
Source of Variation	*SS*	*Df*	*MS*	*F*	*P-value*	*F crit*
Sample	1000.536	1	1000.536	22.531	0.000	4.747
Columns	685.738	2	342.869	7.721	0.007	3.885
Interaction	80.458	2	40.229	0.906	0.430	3.885
Within	532.887	12	44.407			
Total	2299.618	17				

Test statistic: $F_{calc} = 0.906$
Decision: Since $F_{calc} = 0.906$ is less than the critical bound of 3.885, do not reject H_0. There is insufficient evidence to conclude there is an interaction between department and store.

(b) H_0: $\mu_{\text{Sports}} = \mu_{\text{Perfume}}$
H_1: $\mu_{\text{Sports}} \neq \mu_{\text{Perfume}}$
Decision rule: If $F_{calc} > 4.747$, reject H_0.
Test statistic: $F_{calc} = 22.531$
Decision: Since $F_{calc} = 22.531$ is greater than the critical bound of 4.747, reject H_0. There is sufficient evidence to conclude that mean customer service rating does differ between departments.

(c) H_0: $\mu_1 = \mu_2 = \mu_3$
H_1: At least one of the means differ.
Decision rule: If $F_{calc} > 3.885$, reject H_0.
Test statistic: $F_{calc} = 7.721$
Decision: Since $F_{calc} = 7.721$ is greater than the critical bound of 3.885, reject H_0. There is enough evidence to conclude that mean customer service rating does differ between three stores.

(d)



(e) The Tukey procedure should not be used since the null hypothesis in part (a) above was not rejected.

(f) The mean customer service rating is higher for the sports department and it is the highest for store 3.

(g) In 11.51 part (d) the conclusion was that the level of customer service across three stores is about the same. However, in this question, the two-factor experiment gave a more complete, refined set of results than the one-factor experiment.

11.54 (a) To test the homogeneity of variance, you perform a Levene test:

$H_0: \sigma_1^2 = \sigma_2^2 = \sigma_3^2$ $\quad H_1$: Not all σ_j^2 are the same.

ANOVA: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Boost juice	5	7.8	1.56	0.908
Naked juice	5	3.4	0.68	0.232
Funky juice	5	3.9	0.78	0.837

ANOVA

Source of Variation	*SS*	*Df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	2.321333	2	1.160667	1.761254	0.21346	3.885294
Within Groups	7.908	12	0.659			
Total	10.22933	14				

Decision rule: if $F_{calc} > 3.885$ reject H_0
$F_{calc} = 1.76 < 3.885$, do not reject H_0. There is not enough evidence to suggest a significant difference in variances for the three outlets.

(b) $H_0 : \mu_1 = \mu_2 = \mu_3$ H_1: At least one of the means differ.

ANOVA: Single Factor

SUMMARY

Groups	*Count*	*Sum*	*Average*	*Variance*
Boost juice	5	35.6	7.12	3.852
Naked juice	5	32.5	6.5	0.76
Funky juice	5	38.5	7.7	1.285

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	3.601333	2	1.800667	0.916059	0.426339	3.885294
Within Groups	23.588	12	1.965667			
Total	27.18933	14				

Decision rule: If $F_{calc} > 3.885$, reject H_0. $F_{calc} = 0.916 < 3.885$, do not reject H_0. There is insufficient evidence to conclude that the mean price differs across juice outlet.

(c) The Tukey procedure should not be used since the null hypothesis in part (b) above was not rejected.

(d) There is inadequate evidence on which to form a conclusion about differences in mean prices across the three juice outlets.

11.55

ANOVA: Two-Factor With Replication

SUMMARY	Boost juice	Naked juice	Funky juice	Total
University				
Count	4	4	4	12
Sum	30.9	36.4	39.3	106.6
Average	7.725	9.1	9.825	8.883333
Variance	0.669167	0.913333	0.255833	1.328788

Shopping centre				
Count	4	4	4	12
Sum	33.2	39.5	34.6	107.3
Average	8.3	9.875	8.65	8.941667
Variance	4.92	1.395833	1.71	2.686288

Street				
Count	4	4	4	12
Sum	22.1	24.5	31.3	77.9
Average	5.525	6.125	7.825	6.491667
Variance	2.369167	1.449167	2.449167	2.74447

Total			
Count	12	12	12
Sum	86.2	100.4	105.2
Average	7.183333	8.366667	8.766667
Variance	3.730606	3.875152	1.938788

ANOVA

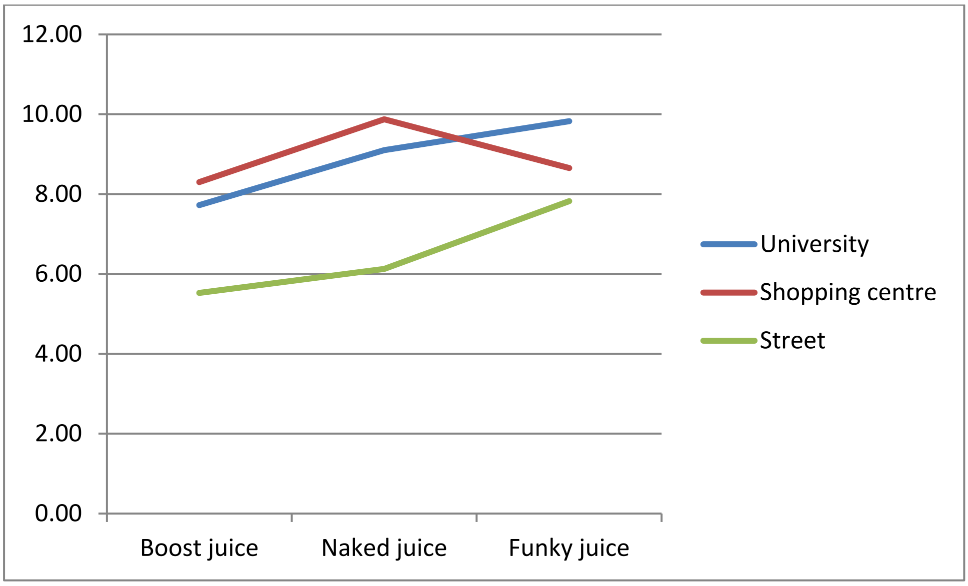
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Sample	46.90389	2	23.45194	13.08405	0.000106	3.354131
Columns	16.26889	2	8.134444	4.538279	0.019993	3.354131
Interaction	9.691111	4	2.422778	1.351689	0.276712	2.727765
Within	48.395	27	1.792407			
Total	121.2589	35				

(a) H_0: There is no interaction between outlet name and location.
H_1: There is an interaction between.
Decision rule: If F_{calc} > 2.7278, reject H_0.
Since F_{calc} = 1.35 < 2.7278, do not reject H_0. There is insufficient evidence to conclude there is an interaction between outlet name and location

(b) $H_0 : \mu_1 = \mu_2 = \mu_3$
H_1: At least one of the means differ.
Decision rule: If F_{calc} > 3.35, reject H_0.
F_{calc} = 4.538 > 3.35, reject H_0.

(c) $H_0 : \mu_1 = \mu_2 = \mu_3$
H_1: At least one of the means differ.
Decision rule: If F_{calc} > 3.35, reject H_0.
F_{calc} = 13.08 > 3.35, reject H_0.

(d)



As there was no interaction effect we may interpret the main effect in (b) and (c)

(e) Mean price differs by outlet and location but there is no interaction between the two.

(f) In the completely randomised design, there is not enough evidence of a difference between mean price for the three outlets. In the two-factor factorial design, there is evidence for differences in mean prices by outlet as well as location. Allowing for the difference in location was necessary to better isolate the difference in price due to outlet.

11.56 (a) Let 1 = more expensive French white, 2 = Italian white, 3 = Italian red, 4 = less expensive French burgundy (red), 5 = more expensive French burgundy (red), 6 = California beaujolais (red), 7 = less expensive French white, and 8 = California white.

H_0: $\mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6 = \mu_7 = \mu_8$

H_1: At least one of the means differs.

Randomised block design output:

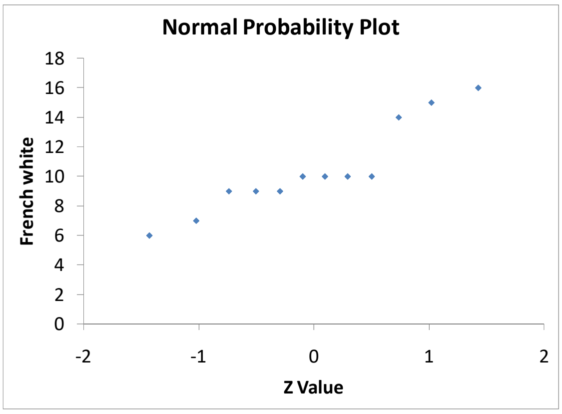
ANOVA

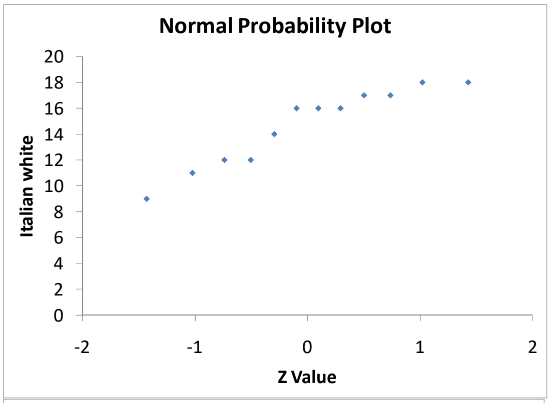
Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Rows	521.5	11	47.409	5.793	0.000	1.915
Columns	440.333	7	62.905	7.686	0.000	2.131
Error	630.167	77	8.184			
Total	1592	95				

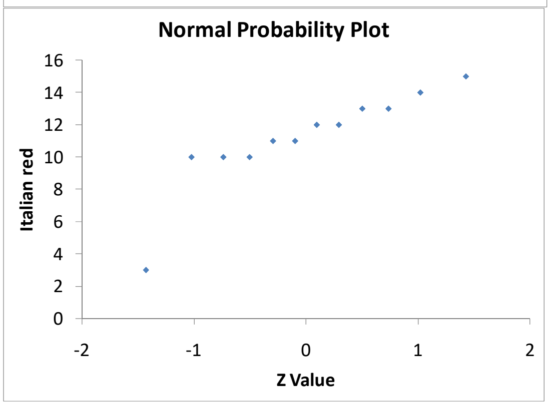
Decision rule: If $F_{calc} > 2.131$, reject H_0.
Test statistic: $F_{calc} = 7.686$
Decision: Since $F_{calc} = 7.686$ is well greater than the critical bound of 2.131, reject H_0.
There is enough evidence of a difference in the mean rating scores among the wines.

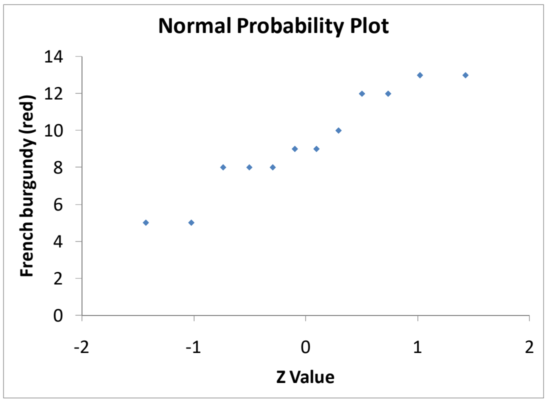
(b) You will need to assume (i) items are randomly selected from the eight groups or are randomly assigned to the eight groups, (ii) the sample values in each group are from a normally distributed population, (iii) the variances of the eight groups are equal, and (iv) there is no interacting effect between the groups and the blocks.

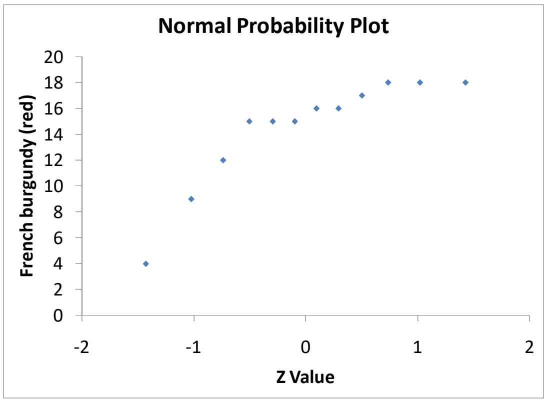
(i) The first assumption is satisfied by the design of the experiment.

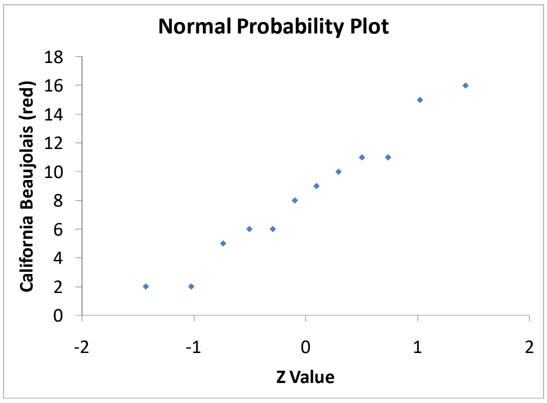


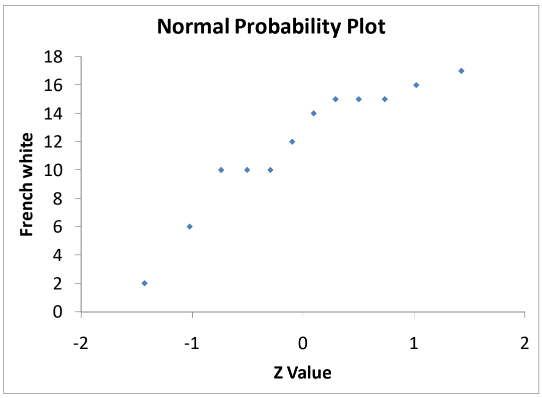


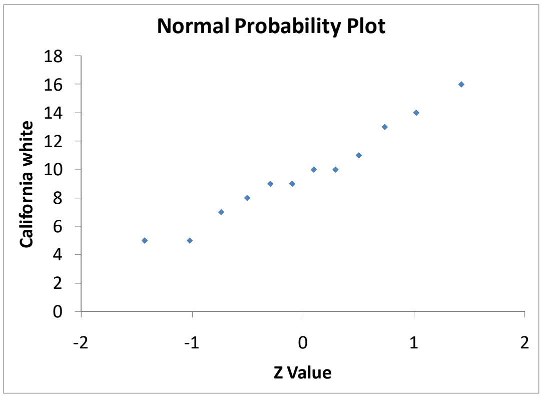



Normal Probability Plot
French burgundy (red)
Z Value
14
12
10
8
6
4
2
0
-2
-1
0
1
2


Normal Probability Plot
French burgundy (red)
Z Value
20
18
16
14
12
10
8
6
4
2
0
-2
-1
0
1
2


Normal Probability Plot
California Beaujolais (red)
Z Value
18
16
14
12
10
8
6
4
2
0
-2
-1
0
1
2


Normal Probability Plot
French white
Z Value
18
16
14
12
10
8
6
4
2
0
-2
-1
0
1
2



With the exception of California beaujolais (red) and California white, the distribution of the other brands shows departure from the normal distribution. However, the randomised block design is quite robust to departure from the normality assumption. Or, alternatively, a Friedman rank sum test can be performed.

(b) ANOVA output for Levene test for homogeneity of variance:

H_0: $\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = \sigma_4^2 = \sigma_5^2 = \sigma_6^2 = \sigma_7^2 = \sigma_8^2$

H_1: At least one of the variances differs.

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	32.167	7	4.595	0.785	0.602	2.116
Within Groups	515.333	88	5.856			
Total	547.5	95				

Since F_{calc} = 0.785<2.116, do not reject the null hypothesis. There is not enough evidence of a difference in the variances.

To test for the blocking effect:

H_0: $\mu_1 = \mu_2 = \ldots = \mu_{12}$

H_1: At least one of the means differs.

Decision rule: If F_{calc} > 1.915, reject H_0.

Test statistic: F_{calc} = 5.793

Decision: Since F_{calc} = 5.793 is greater than the critical bound of 1.915, reject H_0. There is enough evidence to conclude that the mean rating scores differ across the 12 club members. Hence, the blocking has been advantageous in reducing the random error.

(c) To determine which of the wines' mean rating is significantly different from one another, use the Tukey multiple comparisons procedure for randomised block designs. With r = 8 and c = 12, the numerical degrees of freedom is 8 and the denominator degrees of freedom is $(r-1)(c-1) = 7(11)$ = 77. But you will use 8 numerator and 60 denominator degrees of freedom to establish the critical range from Table E.10: Q_α = 4.44

Partial output from the Tukey multiple comparisons procedure for randomised block designs showing only the pairs of groups that are significantly different at the 5% level of significance:

Tukey Multiple Comparisons							
	Sample	**Sample**		Absolute	Std. Error	Critical	
Group	**Mean**	**Size**	Comparison	Difference	of Difference	Range	**Results**
1	10.417	12	Group 1 to Group 2	4.250	0.826	3.667	**Means are different**
2	14.667	12	Group 1 to Group 5	4.000	0.826	3.667	**Means are different**
3	11.167	12	Group 2 to Group 4	5.333	0.826	3.667	**Means are different**
4	9.333	12	Group 2 to Group 6	6.250	0.826	3.667	**Means are different**
5	14.417	12	Group 2 to Group 8	4.917	0.826	3.667	**Means are different**
6	8.417	12	Group 4 to Group 5	5.083	0.826	3.667	**Means are different**
7	11.833	12	Group 5 to Group 6	6.000	0.826	3.667	**Means are different**
8	9.750	12	Group 5 to Group 8	4.667	0.826	3.667	**Means are different**
Other Data							
Level of significance	0.05						
Numerator d.f.	8						
Denominator d.f.	60						
MSE	8.18398						
Q Statistic	4.44						

There are two clusters of wines that have different summated rating. The first group of Italian white, the more expensive French burgundy (red), the less expensive French white and Italian red does not have means that are different from each other significantly within this group while the second group of the more expensive French white, California white, the less expensive French burgundy (red) and California beaujolais (red) does not have means that are different from each other significantly within this group. However, Italian white and the more expensive French burgundy (red) in the first group are significantly rated higher than all the wines in the second group.

(d) From the results in (c), it does not appear that country of origin, the type of wine, or the price has had an effect on the ratings.

(e) $RE = \frac{(r-1)MSBL + r(c-1)MSE}{(rc-1)MSE} = \frac{(11)(47.4091)+12(7)(8.1840)}{(95)(8.1840)}$ = 1.5550

If blocking is not effective, it should be avoided. In this study, there is evidence of a blocking effect, so the completely randomised design model is inferior to the randomised complete block design model. Notice that the relative efficiency measure indicates about 55.5% more observations would be needed in the one-way ANOVA design to obtain the same precision for comparison of treatment group means as would be needed for the randomised complete block design.

11.57 (a)

ANOVA

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Between Groups	440.333	7	62.905	4.807	0.000	2.115
Within Groups	1151.667	88	13.087			
Total	1592	95				

Since F_{calc} = 4.807 > 2.115, reject the null hypothesis. There is enough evidence of a difference in the mean rating scores among the wines.

(b) *SSBL* = 521.5, *SSE* = 630.167, *SSW* = 1151.667. The sum of *SSBL* and *SSE* in the randomised block design equals the *SSW* in the completely randomised design. This has to be true because the *SST*, which measures the total variation of the observations around the grand mean, is the same for both designs and the *SSA*, which measures the difference from group to group, should also be the same for both designs. The *SSW* in the completely randomised design is subdivided into *SSBL* and *SSE* in the randomised block design.

(c) The *p*-value of the randomised block design ($5.0725(10)^{-7}$) is much smaller than that in the completely randomised design (0.000129). Hence, the amount of evidence against the null hypothesis of no treatment effect is much stronger using the randomised block effect, which controls for the variation across the raters.

11.58 Answers may vary.