Exercise 1

Consider comparing weight groups and sexes to see if there is an association between groupings.

a) Construct a contingency table for weight group and sex. Comment on any apparent associations between weight group and sex.

Table of weightgroup by sex				
weightgroup	**sex**			
Frequency Expected	**F**	**I**	**M**	**Total**
heaviest	496 328.86	27 337.67	528 384.47	1051
lightest	106 337	779 346.02	192 393.98	1077
middle	705 641.14	536 658.31	808 749.55	2049
Total	1307	1342	1528	4177

It does appear that there may be a small association between sex and weight group of abalone as there are more than expected heaviest female abalone (Frequency-496, Expected- 328.86), more than expected lightest infant abalone (Frequency-779, Expected- 346.02) and more than expected middle male abalone (Frequency-808, Expected-749.55). Tests will be required to determine if this association is statistically significant.

b) Perform and comment on appropriate tests of association for the table and interpret the results. What do the results tell us about relationships between weight group and sex?

Table of weightgroup by sex				
weightgroup	**sex**			
Frequency Expected	**F**	**I**	**M**	**Total**
heaviest	496 328.86	27 337.67	528 384.47	1051
lightest	106 337	779 346.02	192 393.98	1077
middle	705 641.14	536 658.31	808 749.55	2049
Total	1307	1342	1528	4177

Statistics for Table of weightgroup by sex

Statistic	DF	Value	Prob
Chi-Square	4	1261.6718	<.0001
Likelihood Ratio Chi-Square	4	1384.2153	<.0001
Mantel-Haenszel Chi-Square	1	0.1712	0.6791
Phi Coefficient		0.5496	
Contingency Coefficient		0.4816	
Cramer's V		0.3886	

Sample Size = 4177

It does appear that there may be a small association between sex and weight group of abalone as there are more than expected heaviest female abalone (Frequency-496, Expected- 328.86), more than expected lightest infant abalone (Frequency-779, Expected- 346.02) and more than expected middle male abalone (Frequency-808, Expected-749.55). Tests will be required to determine if this association is statistically significant.

The sample size is large enough for large sample association tests. Pearson’s chi-square and the likelihood ratio chi-square will both be relevant here. Mantel-Haenszel will not be relevant because this test requires that both variables be ordinal. There is no natural ordering for the sex variable and weight group variable.

The Pearson and likelihood ratio statistics are both significant with p-value that is less than 0.001, so the association is statistically significant. The Phi Coefficient is 0.5496, so the association is moderate to strong as indicated by the moderate to strong association from expected in contingency table. The Contingency Coefficient is 0.4816, so the association is moderate as indicated by the moderate association from expected in contingency table. The Cramer’s V is 0.3886, so the association is moderate as indicated by the moderate association from expected in contingency table.

The conclusion is that there is a moderate but statistically significant tendency between sex of abalone and weight group. This implies that there is a probable connection between an abalone's sex and its weight group, and this association is unlikely to be a result of random chance. Since it is not a 2x2 table, we can’t tell they have a positive or negative relationships between two variables.

c) Repeat the analysis but restrict to only the male and female abalone. What can we conclude about weight group differences between male and female abalone combined?

Table of weightgroup by sex			
weightgroup	**sex**		
Frequency Expected	**F**	**M**	**Total**
heaviest	496 472.09	528 551.91	1024
lightest	106 137.38	192 160.62	298
middle	705 697.53	808 815.47	1513
Total	1307	1528	2835

Statistics for Table of weightgroup by sex

Statistic	DF	Value	Prob
Chi-Square	2	15.6982	0.0004
Likelihood Ratio Chi-Square	2	15.9474	0.0003
Mantel-Haenszel Chi-Square	1	0.4433	0.5055
Phi Coefficient		0.0744	
Contingency Coefficient		0.0742	
Cramer's V		0.0744	

Sample Size = 2835

It does appear that there may be a small association between sex and weight group of abalone as there are more than expected heaviest female abalone (Frequency-496, Expected-472.09), more than expected lightest male abalone (Frequency-192, Expected-160.62) and more than expected middle female abalone (Frequency-705, Expected-697.53). Tests will be required to determine if this association is statistically significant.

The sample size is large enough for large sample association tests. Pearson’s chi-square and the likelihood ratio chi-square will both be relevant here. Mantel-Haenszel will not be relevant because this test requires that both variables be ordinal. There is no natural ordering for the sex variable and weight group variable.

The Pearson is significant with p-value that is 0.0004 and likelihood ratio statistics is significant with p-value that is 0.0003, so the association is statistically significant. The related association measures are slightly over 0.07, so the association is weak as indicated by the relatively small deviations from expected in the contingency table.

The conclusion is that there is a small but statistically significant tendency for sex group and weight group. Nonetheless, the degree of this correlation is mild, indicating that although there might be a connection between sex and weight group, it does not significantly influence the determination of an abalone's weight group.

Exercise 2
Now we want to consider the relationship between age groups and weight groups.

a) Construct a contingency table for weight group and age group. Comment on any apparent associations between weight group and age group. Perform and comment on appropriate tests of association for the table and interpret what the results tell us about the relationship between weight group and age group.

Table of weightgroup by agegroup				
weightgroup	**agegroup**			
Frequency **Expected**	**middle**	**oldest**	**youngest**	**Total**
heaviest	398 475.81	648 364.09	5 211.11	1051
lightest	318 487.58	77 373.1	682 216.33	1077
middle	1175 927.62	722 709.82	152 411.57	2049
Total	1891	1447	839	4177

Statistics for Table of weightgroup by agegroup

Statistic	DF	Value	Prob
Chi-Square	4	1961.6008	<.0001
Likelihood Ratio Chi-Square	4	1896.2377	<.0001
Mantel-Haenszel Chi-Square	1	84.8801	<.0001
Phi Coefficient		0.6853	
Contingency Coefficient		0.5653	
Cramer's V		0.4846	

Sample Size = 4177

It does appear that there may be a small association between age group and weight group of abalone as there are less than expected heaviest weight middle age abalone, less than expected lightest weight oldest abalone and less than expected middle weight youngest abalone. Tests will be required to determine if this association is statistically significant.

The sample size is large enough for large sample association tests. Pearson’s chi-square and the likelihood ratio chi-square will both be relevant here. Mantel-Haenszel will not be relevant because this test requires that both variables be ordinal. There is no natural ordering for the age group variable and weight group variable.

The Pearson and likelihood ratio statistics are both significant with p-value that is less than 0.001, so the association is statistically significant. The Phi Coefficient is 0.6583, so the association is strong as indicated by the moderate to strong association from expected in contingency table. The Contingency Coefficient is 0.5653, so the association is moderate to strong as indicated by the moderate to strong association from expected in contingency table. The Cramer’s V is 0.4846, so the association is moderate to strong as indicated by the moderate to strong to association from expected in contingency table.

The conclusion is that there is a moderate to strong but statistically significant tendency for one age group compared to a weight group. Since it is not a 2x2 table, we can’t tell they have a positive or negative relationships between two variables.

b) A researcher is interested in more developed abalone and wants to focus on abalone in the middle and oldest age groups and the middle and heaviest weight groups. Construct the relevant 2x2 contingency table and test for associations between age group and weight group for this subset. Interpret the results and compare them to your findings in part a.

Table of weightgroup by agegroup			
weightgroup	agegroup		
Frequency Expected	middle	oldest	Total
heaviest	398 559.08	648 486.92	1046
middle	1175 1013.9	722 883.08	1897
Total	1573	1370	2943

Statistics for Table of weightgroup by agegroup

Statistic	DF	Value	Prob
Chi-Square	1	154.6604	<.0001
Likelihood Ratio Chi-Square	1	155.5481	<.0001
Continuity Adj. Chi-Square	1	153.7018	<.0001
Mantel-Haenszel Chi-Square	1	154.6079	<.0001
Phi Coefficient		-0.2292	
Contingency Coefficient		0.2234	
Cramer's V		-0.2292	

Fisher's Exact Test	
Cell (1,1) Frequency (F)	398
Left-sided Pr <= F	<.0001
Right-sided Pr >= F	1.0000
Table Probability (P)	<.0001
Two-sided Pr <= P	<.0001

Sample Size = 2943

It does appear that there may be a small association between age group and weight group of abalone as there are less than expected heaviest weight middle age abalone and less than expected middle weight oldest age abalone. Tests will be required to determine if this association is statistically significant.

The sample size is large enough for large sample association tests. Pearson's chi-square and the likelihood ratio chi-square will both be relevant here. Mantel-Haenszel will not be relevant because this test requires that both variables be ordinal. There is no natural ordering for the age group variable and weight group variable.

The Pearson and likelihood ratio statistics are both significant with p-value that is less than 0.001, so the association is statistically significant. The Phi Coefficient is -0.2292, so the association is moderate as indicated by the moderate from expected in contingency table. The Contingency Coefficient is 0.2234, so the association is moderate as indicated by the moderate association from expected in contingency table. The Cramer's V is -0.2292, so the association is moderate as indicated by the moderate association from expected in contingency table.

Fisher's exact test is not needed here, but a comparison of the chi-square p-values with the two-sided Fisher p-value shows how good the asymptotic approximations are in this case with a sample size of just over 500. The approximate tests have p-values is less than 0.0001 while the exact test has a p-value also less than 0.0001.

The conclusion is that there is a moderate negative but statistically significant tendency for one age group of abalone to be less likely to be in a weight group and the other has a slightly lower tendency to be in other weight group. A look at the data shows that heaviest are little less often in middle group than expected due to chance, middle weight is less like in middle age group.

For part(a), we draw a conclusion saying that the two variables have moderate to strong association, but we don't know positive or negative relationships between two variables because it is not a 2x2 table. But in part(b), since we know that the table is a 2x2 table, we know that that is a negative relationship and statistically significant association between two variables since Phi Coefficient value and Cramer's V values are negative.

c) For the 2x2 table, test whether the oldest abalone are at higher risk than the middle age abalone to be in the heaviest weight group and state your conclusions.

Table of agegroup by weightgroup			
agegroup	**weightgroup**		
Frequency Row Pct	**heaviest**	**middle**	**Total**
middle	398 25.30	1175 74.70	1573
oldest	648 47.30	722 52.70	1370
Total	1046	1897	2943

Statistics for Table of agegroup by weightgroup

Column 1 Risk Estimates						
	Risk	**ASE**	**95% Confidence Limits**		**Exact 95% Confidence Limits**	
Row 1	0.2530	0.0110	0.2315	0.2745	0.2317	0.2753
Row 2	0.4730	0.0135	0.4466	0.4994	0.4463	0.4998
Total	0.3554	0.0088	0.3381	0.3727	0.3381	0.3730
Difference	-0.2200	0.0174	-0.2540	-0.1859		
Difference is (Row 1 - Row 2)						

Sample Size = 2943

Column risk estimate tables are provided for both columns of the contingency table. The column 1 risk for row 1 is estimated to be .2530 based on the sample. This is the estimate of the risk of being in the heaviest weight group given that an abalone is middle age group. The estimate of the risk of being in the heaviest weight group given that an abalone is oldest age group is .4730.

Since the question is asking the oldest abalone are at higher risk than the middle age abalone in the heaviest weight group so column risks for the first column can be used. The estimated difference is -0.220 with an asymptotic confidence interval of (-0.2540, -0.1859). The difference is moderate, and the entire confidence interval is negative, so the risk for oldest abalone being in the heaviest weight group is statistically significantly higher. The approximate confidence intervals are very close to the exact intervals in this case because the sample sizes are somewhat large for both middle and oldest of abalone.

Exercise 3
The data contains the actual weights, so differences can be tested and quantified across groups in this case. Assume
the normality assumption is reasonable, so you can proceed without testing normality.
a) Perform a one-way ANOVA for whole weight with sex as the categorical predictor. Test any assumptions of the model that should be tested (aside from normality, which you do not need to test). Comment on the significance of the model and the variation described by the model.

Class Level Information		
Class	**Levels**	**Values**
sex	3	F I M

Number of Observations Read	4177
Number of Observations Used	4177

Dependent Variable: whole_weight

Some basic information about the data is provided first, namely the classification variable(s), levels and values are 3 for the classification variable(s) (F, I, M) and the amount of data read and used which is 4177. Then analysis of variance tables and diagnostics follow. The number of Observations read and used are the same meaning that we don’t lose any data.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	314.366955	157.183478	951.01	<.0001
Error	4174	689.883326	0.165281		
Corrected Total	4176	1004.250281			

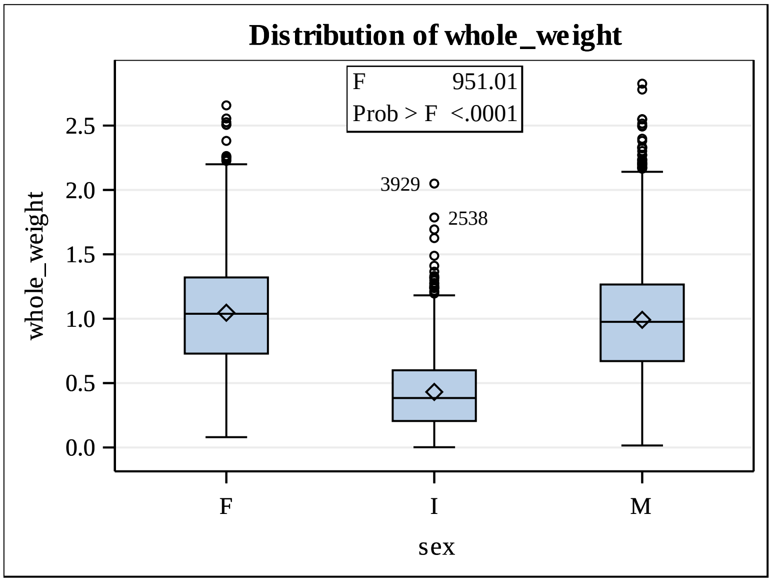
In this example, the p-value is less than 0.0001, so the null model is rejected in favor of the one-way sex model. The variation described by differences across sex is significantly greater (from a statistical perspective) than expected due to chance.

R-Square	Coeff Var	Root MSE	whole_weight Mean
0.313036	49.05600	0.406548	0.828742

Looking at the R-Square value in the table that follows would provide some indication of practical significance. The R-Square value, which can be obtained by dividing the sum of squares for the model by the total sum of squares, tells the percentage of overall variation in the response described by the model. In this example, about 31.3% of the variation in whole_weight could be described by sex alone. This is a moderate percentage, and it is a good start, adding additional explanatory variables to the model will help to describe more of the variation in whole weight values.

Source	DF	Anova SS	Mean Square	F Value	Pr > F
sex	2	314.3669553	157.1834776	951.01	<.0001

It is worth noting that R-Square values will not go down if additional terms are added to a model. The amount of increase may be too small to be of practical value, though. The table is referred to as the model ANOVA table in SAS. It breaks down the model into specific terms and their contributions to the overall variation explained. In this case, there is only one term, so the sum of squares for the overall model is the same as the sum of squares for the sex term.



Under the null hypothesis, the means would be the same for each of the sex and the boxes would line up horizontally. The shifts in the box plots provide some indications of possible significantly different means. Additional testing will be needed to determine which differences are statistically significant and to quantify those differences.

Levene's Test for Homogeneity of whole_weight Variance ANOVA of Squared Deviations from Group Means					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
sex	2	14.6370	7.3185	102.21	<.0001
Error	4174	298.9	0.0716		

Welch's ANOVA for whole_weight			
Source	DF	F Value	Pr > F
sex	2.0000	1307.36	<.0001
Error	2667.9		

For Levene's test, the null hypothesis is that the variances are the same for each group (each sex in this case). Just like with the Folded F test, Levene's test is one-sided. A large F value and a corresponding small p-value would indicate some of the variances are far from equal and adjustments should be made.

In the sex model, the p-value is less than 0.0001, so we reject the null hypothesis, and we know that it is a different variance. Levene's test had been significant, the following Welch's ANOVA table would be the basis for deciding whether the overall model is better than an error-only model.

The F test is interpreted in a similar way to that for the usual overall ANOVA table—a significant F statistic indicates the systematic part of the model describes significantly more variation in the response than expected due to chance, and that the model is better than the null error-only model.

The conclusion here is that the sex term describes more variation than expected due to chance and it would be a good idea to keep that term in the model since we know that from Welch's ANOVA model. We reject the null hypothesis which is an error-only model and favor the model with sex variables. We also find out that 31.3% of the variation in whole_weight could be described by sex alone.

b) Comment on any significantly different whole weight means as determined by the best test for comparing all pairwise differences. Explain what that tells us about differences in whole weight across the sexes, and comment on how these results compare with the results from Exercise 1

Tukey's Studentized Range (HSD) Test for whole_weight

Note: This test controls the Type I experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	4174
Error Mean Square	0.165281
Critical Value of Studentized Range	3.31568

A table containing alpha value which is 0.05, degrees of freedom which is 4174, and a few statistics needed for constructing confidence intervals for the differences of means is generated and followed by the comparison of means.

Comparisons significant at the 0.05 level are indicated by ***.				
sex Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		
F - M	0.05507	0.01916	0.09099	***
F - I	0.61517	0.57813	0.65221	***
M - F	-0.05507	-0.09099	-0.01916	***
M - I	0.56010	0.52444	0.59576	***

I - F	-0.61517	-0.65221	-0.57813	***
I - M	-0.56010	-0.59576	-0.52444	***

The estimated differences of means are provided along with 95% confidence intervals that have considered the multiple comparisons being made. The table shows twice as many comparisons as needed because it includes both a-b and b-a for each possible value of a and b in the data. The final column contains indicators for significance at a 0.05 level. Each row that contains *** is a statistically significant difference.

From the results, it can be concluded that mean whole weight for infant is significantly lower than for females and males. All the comparison about sex variables to whole weight variables in the table is statistically significant. Further, female abalones are expected to have mean whole weight 0.05507 higher than male abalones on average which tell us that their difference of means are small, and the 95% confidence interval would be (0.01916, 0.09099). Male abalones are expected to have whole weight on average 0.56 higher than those with infant abalone, with a confidence interval of (0.52444, 0.59576).

The fact that 0 is in neither confidence interval demonstrates the significant difference—a difference of 0 would be no difference, and 0 is not in the confidence intervals for these two comparisons.

Problem 1 shows that there is a moderate but statistically significant tendency between sex of abalone and weight group for male, female and infant’s abalone. This implies that there is a probable connection between an abalone's sex and its weight group, and this association is unlikely to be a result of random chance. Problem 1 also shows that there is a small but statistically significant tendency for sex group and weight group for only male and female abalone. Nonetheless, the degree of this correlation is mild, indicating that although there might be a connection between sex and weight group, it does not significantly influence the determination of an abalone's weight group.

The difference between Problem 1 and Problem 3 is they used different test. Problem 1 using Chi-Square test to check the association between two variables. Problem 3 using Tukey’s method makes its adjustment under the assumption that all pairwise for comparisons will be made and testing. Problem 3 testing all pairwise has different whole weight means and construct a 95% confidence interval.

We can conclude that the association between sex and weight from Problem 1 is reflected in the actual differences in weight means between sexes in Problem 3. This supports the association identified between sex and weight in Problem 1, providing empirical evidence that there are distinct weight characteristics associated with different sexes.

Exercise 4

a) Perform a one-way ANOVA for whole_weight with agegroup as the categorical predictor. Test any assumptions of the model that should be tested (aside from normality, which you do not need to test). Comment on the significance of the model and the variation described by the model.

Class Level Information		
Class	**Levels**	**Values**
agegroup	3	middle oldest youngest

Number of Observations Read	4177
Number of Observations Used	4177

Dependent Variable: whole_weight

Some basic information about the data is provided first, namely the classification variable(s), levels and values are 3 for the classification variable(s) (middle, oldest, youngest) and the amount of data read and used which is 4177. Then analysis of variance tables and diagnostics follow. The number of Observations read and used are the same meaning that we don't lose any data.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	368.946114	184.473057	1212.00	<.0001
Error	4174	635.304167	0.152205		
Corrected Total	4176	1004.250281			

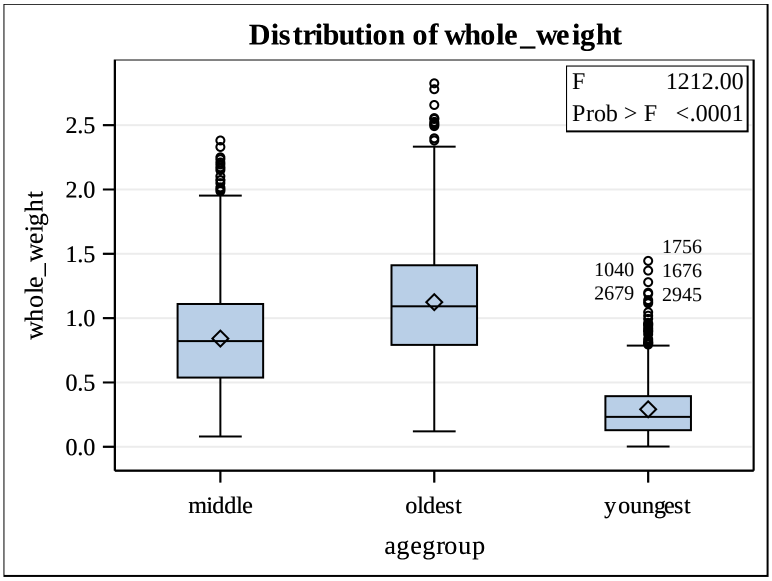
In this example, the p-value is less than 0.0001, so the null model is rejected in favor of the one-way sex model. The variation described by differences across age is significantly greater (from a statistical perspective) than expected due to chance.

R-Square	Coeff Var	Root MSE	whole_weight Mean
0.367385	47.07553	0.390135	0.828742

Looking at the R-Square value in the table that follows would provide some indication of practical significance. The R-Square value, which can be obtained by dividing the sum of squares for the model by the total sum of squares, tells the percentage of overall variation in the response described by the model. In this example, about 36.7% of the variation in whole_weight could be described by age alone. This is a moderate percentage, and it is a good start, adding additional explanatory variables to the model will help to describe more of the variation in whole weight values.

Source	DF	Anova SS	Mean Square	F Value	Pr > F
agegroup	2	368.9461142	184.4730571	1212.00	<.0001

It is worth noting that R-Square values will not go down if additional terms are added to a model. The amount of increase may be too small to be of practical value, though. The table is referred to as the model ANOVA table in SAS. It breaks down the model into specific terms and their contributions to the overall variation explained. In this case, there is only one term, so the sum of squares for the overall model is the same as the sum of squares for the age term.



Under the null hypothesis, the means would be the same for each of the sex and the boxes would line up horizontally. The shifts in the box plots provide some indications of possible significantly different means. Additional testing will be needed to determine which differences are statistically significant and to quantify those differences.

Levene's Test for Homogeneity of whole_weight Variance ANOVA of Squared Deviations from Group Means					
Source	**DF**	**Sum of Squares**	**Mean Square**	**F Value**	**Pr > F**
agegroup	2	13.9608	6.9804	120.10	<.0001
Error	4174	242.6	0.0581		

Welch's ANOVA for whole_weight			
Source	**DF**	**F Value**	**Pr > F**
agegroup	2.0000	2106.56	<.0001
Error	2620.4		

For Levene’s test, the null hypothesis is that the variances are the same for each group (each sex in this case). Just like with the Folded F test, Levene’s test is one-sided. A large F value and a corresponding small p-value would indicate some of the variances are far from equal and adjustments should be made.

In the sex model, the p-value is less than 0.0001, so we reject the null hypothesis, and we know that it is a different variance. Levene’s test had been significant, the following Welch’s ANOVA table would be the basis for deciding whether the overall model is better than an error-only model.

The F test is interpreted in a similar way to that for the usual overall ANOVA table—a significant F statistic indicates the systematic part of the model describes significantly more variation in the response than expected due to chance, and that the model is better than the null error-only model.

The conclusion here is that the age term describes more variation than expected due to chance and it would be a good idea to keep that term in the model since we know that from Welch's ANOVA model. We reject the null hypothesis which is an error-only model and favor the model with age variables. We also find out that 36.7% of the variation in whole_weight could be described by age alone.

b) Comment on any significantly different whole weight means as determined by the best test for comparing all pairwise differences. Explain what that tells us about differences in whole weight across age groups, and comment on how these results compare with the results from Exercise 2.

Tukey's Studentized Range (HSD) Test for whole_weight

Note: This test controls the Type I experimentwise error rate.

Alpha	0.05
Error Degrees of Freedom	4174
Error Mean Square	0.152205
Critical Value of Studentized Range	3.31568

A table containing alpha value which is 0.05, degrees of freedom which is 4174, and a few statistics needed for constructing confidence intervals for the differences of means is generated and followed by the comparison of means.

Comparisons significant at the 0.05 level are indicated by ***.				
agegroup Comparison	**Difference Between Means**	**Simultaneous 95% Confidence Limits**		
oldest - middle	0.28310	0.25115	0.31505	***
oldest - youngest	0.83290	0.79321	0.87259	***
middle - oldest	-0.28310	-0.31505	-0.25115	***
middle - youngest	0.54980	0.51186	0.58775	***
youngest - oldest	-0.83290	-0.87259	-0.79321	***
youngest - middle	-0.54980	-0.58775	-0.51186	***

The estimated differences of means are provided along with 95% confidence intervals that have considered the multiple comparisons being made. The table shows twice as many comparisons as needed because it includes both a-b and b-a for each possible value of a and b in the data. The final column contains indicators for significance at a 0.05 level. Each row that contains *** is a statistically significant difference.

From the results, it can be concluded that mean whole weight for youngest age group is significantly lower than for middle age group and oldest age group. All the comparison about age variables to whole weight variables in the table is statistically significant. Further, oldest abalones are expected to have mean whole weight 0.28310 higher than middle size of abalones on average which tell us that their difference of means are small, and the 95% confidence interval would be (0.25115,0.31505). Middle size of abalones are expected to have whole weight on average 0.54980 higher than those with youngest age abalone, with a confidence interval of (0.51186, 0.58775).

The fact that 0 is in neither confidence interval demonstrates the significant difference—a difference of 0 would be no difference, and 0 is not in the confidence intervals for these two comparisons.

For Problem 2, the conclusion is that there is a moderate negative but statistically significant tendency for one age group of abalone to be less likely to be in a weight group and the other has a slightly lower tendency to be in other weight group. A look at the data shows that heaviest are little less often in middle group than expected due to chance, middle weight is less like in middle age group. For Problem2, without youngest and lightest variable, the conclusion is that there is a moderate to strong but statistically significant tendency for one age group compared to a weight group.

The difference between Problem 2 and Problem 4 is they used different test. Problem 2 using Chi-Square test to check the association between two variables. Problem 4 using Tukey's method makes its adjustment under the assumption that all pairwise for comparisons will be made and testing. Problem 4 testing all pairwise age variable has different whole weight means and construct a 95% confidence interval.

We can conclude that the association between age and weight from Problem 2 is reflected in the actual differences in weight means between the three group of ages in Problem 4. This supports the association identified between age and weight in Problem 2, providing empirical evidence that there are distinct weight characteristics associated with different ages.