Worksheet 4 - Analysis of variance (ANOVA)

Analysis of variance (ANOVA) references

- Fowler *et al.* (1998) -Chpt 17
- Holmes *et al.* (2006) - Chpt 7
- Quinn & Keough (2002) - Chpts 8 & 9 (mainly Section 9.2)

Question 1 - ANOVA and Tukey's test

Here is a modified example from Quinn and Keough (2002). Day and Quinn (1989) described an experiment that examined how rock surface type affected the recruitment of barnacles to a rocky shore. The experiment had a single factor, surface type, with 4 treatments or levels: algal species 1 (ALG1), algal species 2 (ALG2), naturally bare surfaces (NB) and artificially scraped bare surfaces (S). There were 5 replicate plots for each surface type and the response (dependent) variable was the number of newly recruited barnacles on each plot after 4 weeks.

Format of day.csv data files

TREAT	BARNACLE
ALG1	27
..	..
ALG2	24
..	..
NB	9
..	..
S	12
..	..

TREAT   Categorical listing of surface types. ALG1 = algal species 1, ALG2 = algal species 2, NB = naturally bare surface, S = scraped bare surface.

BARNACLE   The number of newly recruited barnacles on each plot after 4 weeks.
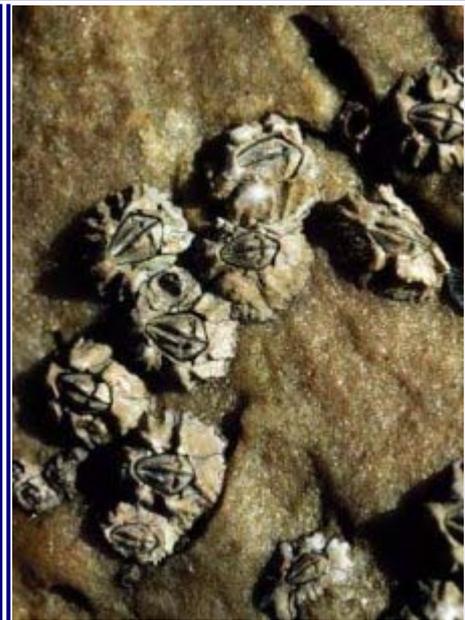
Open the day data file.

Note that as with independent t-tests, variables are in columns with levels of the categorical variable listed repeatedly. Day and Quinn (1989) were interested in whether substrate type influenced barnacle recruitment. This is a biological question. To address this question statistically, it is first necessary to re-express the question from a statistical perspective.

Q1-1. From a classical hypothesis testing point of view, what is the statistical question they are investigating? That is, what is their statistical H_0?

Q1-2. The appropriate statistical test for comparing the means of more than two groups, is an ANOVA. In the table below, list the assumptions of **ANOVA** along with how violations of each assumption are diagnosed and/or the risks of violations are minimized.

Assumption	Diagnostic/Risk Minimization
I.	
II.	
III.	

Using **boxplots**, examine the assumptions of normality and homogeneity of variance. Note that when sample sizes are small (as is the case with this data set), these ANOVA assumptions cannot reliably be checked using boxplots since boxplots require at least 5 replicates (and preferably more), from which to calculate the median and quartiles. As with regression analysis, it is the assumption of homogeneity of variances (and in particular, whether there is a relationship between the mean and variance) that is of most concern for ANOVA.

Q1-3. Check the assumption of homogeneity of variances by **plotting the sample (group) means against sample variances**. A strong relationship (positive or negative) between mean and variance suggests that the assumption of homogeneity of variances is likely to be violated.

    a. Any evidence of non-homogeneity? (Y or N)

Q1-4. Test the null hypothesis that the population group means are equal using a **single factor ANOVA**. As with regression analysis, it is also a good habit to **examine the resulting diagnostics** (note that Leverage and thus Cook's D have no useful meaning for categorical X variables) and residual plot. If there are no obvious problems, then the analysis is likely to be reliable. **Examine the ANOVA table**.

Q1-5. **Identify** the important items from the ANOVA output and fill out the following ANOVA table

Source of Variation	SS	df	MS	F-ratio
Between groups				
Residual (within groups)				

Q1-6. What is the probability that the observed samples (and the degree of differences in barnacle recruitment between them) or ones more extreme, could be collected from populations in which there are no differences in barnacle recruitment. That is, what is the probability of having the above F-ratio or one more extreme when the null hypothesis is true?

Q1-7. What statistical conclusion would you draw?

Q1-8. Write the results out as though you were writing a research paper/thesis. For example (select the phrase that applies and fill in gaps with your results):
The mean number of barnacles recruiting was (choose the correct option)

(choose correct option   (F =   , df =   ,   , P =   )
different from the mean metabolic rate of female fulmars.

Q1-9. Such a table could be incorporated into the results section of a report. **Copy and paste the ANOVA table from the Rcmdr output window into Word, format the table correctly and add an**

appropriate table caption.

Q1-10. Such a result would normally be accompanied by a graph to illustrate the mean (and variability or precision thereof) barnacle recruitment on each substrate type. Construct such a **bar graph** showing the mean barnacle recruitment on each substrate type and an indication of the precision of the means with error bars. To see how these results could be incorporated into a report, **save the graph (as a jpeg)** and **import the picture into Word**.

Q1-11. Although we have now established that there is a statistical difference between the group means, we do not yet know which group(s) are different from which other(s). For this data a **Tukey's multiple comparison test** (to determine which surface type groups are different from each other, in terms of number of barnacles) is appropriate. **Complete the following table for Tukey's pairwise comparison of group means:** include differences between group means (**ignore the sign**) and Tukey's adjusted P-values (in brackets) for each pairwise comparison.

	ALG1	ALG2	NB	S
ALG1	0.000 (1.00)			
ALG2	6.000 (0.165)	0.000 (1.000)		
NB	☐ (☐)	☐ (☐)	0.000 (1.000)	
S	☐ (☐)	☐ (☐)	☐ (☐)	0.000 (1.000)

Q1-12. What are your conclusions from the Tukey's tests?

Q1-13.A way of representing/summarizing the results of multiple comparison tests is to **incorporate symbols** into the bargraph such that similarities and differences in the symbols associated with bars reflect statistical outcomes. Produce such a graph, **save the graph as a picture** and **import the picture into Microsoft Word**. Make sure you also incorporate an appropriate figure caption under the graph.

Question 2 - ANOVA and Tukey's test

Here is a modified example from Quinn and Keough (2002). Medley & Clements (1998) studied the response of diatom communities to heavy metals, especially zinc, in streams in the Rocky Mountain region of Colorado, U.S.A.. As part of their study, they sampled a number of stations (between four and seven) on six streams known to be polluted by heavy metals. At each station, they recorded a range of physiochemical variables (pH, dissolved oxygen etc.), zinc concentration, and variables describing the diatom community (species richness, species diversity H and proportion of diatom cells that were the early-successional species, *Achanthes minutissima*). One of their analyses was to ignore streams and partition the 34 stations into four zinc-level categories: background (< 20 µg.l^{-1}, 8 stations), low (21-50 µg.l^{-1}, 8 stations), medium (51-200 µg.l^{-1}, 9 stations), and high (> 200 µg.l^{-1}, 9 stations) and test null hypotheses that there we no differences in diatom species diversity between zinc-level groups, using stations as replicates. We will also use these data to test the null hypotheses that there are no differences in diatom species diversity between streams, again using stations as replicates.

Format of medley.csv data files

STATION	ZINC	DIVERSITY
ER1	BACK	2.27
...	...	...
ER2	HIGH	1.25
...	...	...

EF1	LOW	1.4
...	...	...
ER4	MEDIUM	1.62
...	...	...

STATION	Uniquely identifies the sampling station from which the data were collected.
ZINC	Zinc level concentration categories.
DIVERSITY	Shannon-Weiner species diversity of diatoms
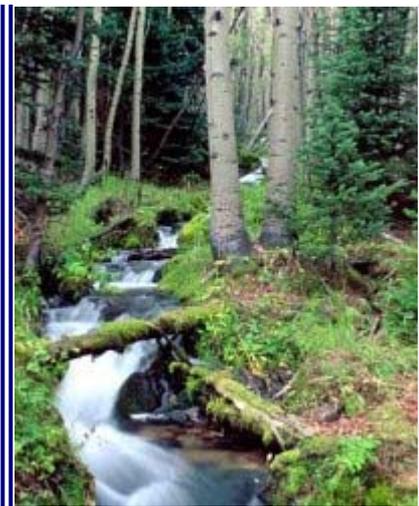
Open the medley data file.

Most statistical packages automatically order the levels of categorical variables alphabetically. Therefore, the levels of the ZINC categorical variable will automatically be ordered as (BACK, HIGH, LOW, MEDIUM). For some data sets the ordering of factors is not important. However, in the medley data set, it would make more sense if the factors were in the following order (BACK, LOW, MEDIUM, HIGH) as this would more correctly represent the relationships between the levels. Note that the ordering of a factor has no bearing on any analyses, it just makes the arrangement of data summaries within some graphs and tables more logical. It is therefore recommended that whenever a data set includes categorical variables, **reorder the levels of these variables** into a logical order.

Q2-1. Check the ANOVA assumptions using a **factorial boxplot**.

   a. Any evidence of skewness? ☐

   b. Outliers? ☐

   c. Does the spread of data look homogeneous between the different Zinc levels? ☐

If the assumptions seem reasonable, **fit the linear model**, check the **residuals** and if still there is no clear indication of problems, **examine the ANOVA table**.

Q2-2. Write the results out as though you were writing a research paper/thesis. For example (select the phrase that applies and fill in gaps with your results):
The mean diatom diversity was (choose the correct option)
(choose correct option ☐ (F = ☐, df = ☐, ☐, P = ☐)
different between the four zinc level groups.

This can be abbreviated to $F_{df_{Groups}, df_{Residual}}$=fratio, P=pvalue. To see how the full anova table might be included in a report/thesis.

Q2-3. Copy and paste the ANOVA table from the Rcmdr output window into Word and add an appropriate table caption .

Q2-4. Now determine which zinc level groups were different from each other, in terms of diatom species diversity, using a **Tukey's multiple comparison test**. **Incorporate symbols** these findings onto a **bargraph**. Note, it is important that you recall the order of the factors. Note also that some groups may need multiple symbols. **Saving the graph as a jpeg image** and **import the graph into word**.

Question 3 - ANOVA and planned comparisons

Here is a modified example from Quinn and Keough (2002). Partridge and Farquhar (1981) set up an experiment to examine the effect of reproductive activity on longevity (response variable) of male fruitflies (Drosophila sp.). A total of 125 male fruitflies were individually caged and randomly assigned to one of five treatment groups. Two of the groups were used to to investigate the effects of the number of partners (potential matings) on male longevity, and thus differed in the number of female partners in the cages (8 vs 1). There were two corresponding control groups containing eight and one newly pregnant female partners (with which the male flies cannot mate), which served as controls for any potential effects of competition between males and females for food or space on male longevity. The final group had no partners, and was used as an overall control to determine the longevity of un-partnered male fruitflies.

Format of partridge.csv data files

GROUP	LONGEVITY
PREG8	35
..	..
NON0	40
..	..
PREG1	46
..	..
VIRGIN1	21
..	..
VIRGIN8	16
..	..

GROUP   Categorical listing of female partner type.
PREG1 = 1 pregnant partner, NONE0 = no female partners, PREG8 = 8 pregnant partners, VIRGIN1 = 1 virgin partner, VIRGIN8 = 8 virgin partners.
Groups 1,2,3 - Control groups
Groups 4,5 - Experimental groups.

LONGEVITY   Longevity of male fruitflies (days)
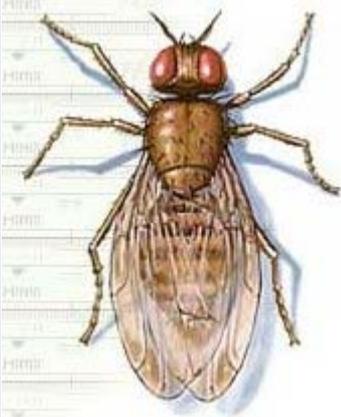
Open the partridge data file.

Q3-1. When comparing the mean male longevity of each group, what is the null hypothesis?

Note, normally we might like to adjust the ordering of the levels of the categorical variable (GROUP), however, in this case, the alphabetical ordering also results in the most logical ordering of levels.

Q3-2. Before performing the ANOVA, check the assumptions (Boxplots, scatterplot of Mean vs Variance) using the variable GROUP as the grouping (IV) variable for the X-axes. Is there any evidence that the ANOVA assumptions have been violated (Y or N)? ☐

In addition to the global ANOVA in which the overall effect of the factor on male fruit fly longevity is examined, a number of other comparisons can be performed to identify differences between specific groups. As with the previous question, we could perform Tukey's post-hoc pairwise comparisons to examine the differences between each group. Technically, it is only statistically legal to perform n-1 pairwise comparisons, where n is the number of groups. This is because if each individual comparison excepts a 5% (&alpha=0.05) probability of falsely rejecting the H_0, then the greater the number of tests performed the greater the risk eventually making a statistical error. Post-hoc tests protect against such an outcome by adjusting the &alpha values for each individual comparison down. Consequently, the power of each comparison is greatly reduced.

This particular study was designed with particular comparisons in mind, while other pairwise comparisons would have very little biological meaning or importance. For example, in the context of the project aims, comparing group 1 with group 2 would not yield anything meaningful. As we have five groups (df=4), we can do four **planned comparisons**.

Q3-3. In addition to the global ANOVA, we will address two specific questions by planned comparisons.

   a. "Is longevity affected by the presence of a large number of potential mates (8 virgin females compared to 1 virgin females)?" (contrast coefficients: 0, 0, 0, -1, 1)

   b. "Is longevity affected by the presence of any number of potential mates compared with either no partners or pregnant partners?" (contrast coefficients: -2, -2, -2, 3, 3)

Q3-4. Before we fit the linear model (perform the ANOVA), we need to define the contrast coefficients (and thus comparisons) that we wish to perform in addition to the global ANOVA. **Define the contrasts for the GROUP variable**.

Q3-5. If there is no evidence that the assumptions have been violated and the contrasts were successfully defined, **run the linear model** and **examine the ANOVA table**.

Q3-6. Present the results of the global ANOVA and planned comparisons as part of the following ANOVA table:

Source of Variation	SS	df	MS	F-ratio	Pvalue
Between groups	☐	☐	☐	☐	☐
8 virgin vs 1 virgin	☐	☐	☐	☐	☐
(1 virg + 8 virg) vs (control + preg)	☐	☐	☐	☐	☐
Residual (within groups)	☐	☐	☐		

Note that the Residual (within groups) term is common to each planned comparison as well as the original global ANOVA. **Copy and paste the ANOVA table from the Rcmdr output window into Word and add an appropriate table caption** .

Q3-7. Summarize the conclusions (statistical and biological) from the analyses.

   a. Global null hypothesis (H_0: population group means all equal)

   b. Planned comparison 1 (H_0: population mean of 8VIRGIN group is equal to that of 1VIRGIN)

   c. Planned comparison 2 (H_0: population mean of average of 1VIRGIN and 8VIRGIN groups are equal to the population mean of average of CONTROL, 1PREG and 8PREG groups)

Q3-8. List any other specific comparisons that may have been of interest to this study. Remember that the total number of comparisons should not exceed the global degrees of freedom (4 in this case) and each outcome of each comparison should be independent of all other comparisons.

Q3-9. Finally, construct an appropriate **graph** to accompany the above analyses. **Save the graph as a jpeg image** and **import the graph into word**.

Question 4 - ANOVA and planned comparisons

Snodgrass et al. (2000) were interested in how the assemblages of larval amphibians varied between depression wetlands in South Carolina, USA, with different hydrological regimes. A secondary question was whether the presence of fish, which only occurred in wetlands with long hydroperiods, also affected the assemblages of larval amphibians. They sampled 22 wetlands in 1997 (they originally had 25 but three dried out early in the study) and recorded the species richness and total abundance of larval amphibians as well as the abundance of individual taxa. Wetlands were also classified into three hydroperiods: short (6 wetlands), medium (5) and long (11) - the latter being split into those with fish (5) and those without (6). The short and medium hydroperiod wetlands did not have fish.

The overall question of interest is whether species richness differed between the four groups of wetlands. However, there are also specific questions related separately to hydroperiod and fish. Is there a difference in species richness between long hydroperiod wetlands with fish and those without? Is there a difference between the hydroperiods for wetlands without fish? We can address these questions with a single factor fixed effects ANOVA and planned contrasts using species richness of larval amphibians as the response variable and hydroperiod/fish category as the predictor (grouping variable).

Format of snodgrass.csv data files	

HYDROPERIOD	RICHNESS
SHORT	3
..	..
MEDIUM	9
..	..
LONGNOFISH	7
..	..
LONGFISH	12
..	..

HYDROPERIOD	Categorical listing of the four hydroperiod/fish wetlands (short, medium and longnofish represent the hydroperiods of wetlands without fish; longfish represents wetlands with long hydroperiods that contain fish).
RICHNESS	Species richness of larval amphibians

Open the snodgrass data file.

Reorder the factor levels of HYDROPERIOD into a more logical order (e.g. SHORT, MEDIUM, LONGNOFISH, LONGFISH)

Q4-1. Examine the **group means and variances** and **boxplots** for species richness across the wetland categories. Is there any evidence that any of the assumptions have been violated? ('Y' or 'N')

☐

Q4-2. As well as the overall analysis, Snodgrass et al. (2000) were particularly interested in two specific comparisons a) whether there was a difference in species richness between the long hydroperiod wetlands with and without fish, and b) whether there was a difference in species richness between permanent wetlands (long hydroperiods) and temporary wetlands (short and medium hydroperiods). What specific null hypotheses are being tested;

a.

b.

Q4-3. Define the appropriate contrast coefficients (and thus comparisons). Although it was not necessarily obvious, attempting to define these contrasts indicated that the contrasts were not orthogonal (independent). As a result it is only possible to test one of the researchers null hypotheses. Choice one of them and define the appropriate contrasts.

Q4-4. Now **fit a single factor ANOVA model** and **examine the residuals**. Any evidence of skewness or unequal variances? Any outliers? Any evidence of violations? ('Y' or 'N') ☐

Q4-4. Examine the single factor ANOVA and specific comparisons. Fill in the following table (not that the table includes space to fill out the outcome of the two requested planned comparisons, yet only one is possible - just fill out the one that you decided to test!):

Source of Variation	SS	df	MS	F-ratio	Pvalue
Between groups	☐	☐	☐	☐	☐
Long with vs nofish	☐	☐	☐	☐	☐
Permanent vs Temporary	☐	☐	☐	☐	☐
Residual (within groups)	☐	☐	☐		

Q4-5. What statistical conclusions would you draw from the overall ANOVA and the two specific contrasts and what do they mean biologically?

Q4-6. Finally, construct an appropriate **graph** to accompany the above analyses. **Save the graph as a jpeg image** and **import the graph into word**.

Question 5 - Experimental design

A marine biologist was studying the effects of increased lead concentration ([Pb]) in the water on deformities in sperm cells of male flathead in Port Phillip Bay. The aquarium room at the Queenscliff Marine Station was used. There were two large tanks used for the experiment. In one tank (treatment tank), 50l of seawater was removed weekly and replaced with 50l of seawater with a [Pb] of 1 mg/l. In the other tank, the seawater was not altered (control tank). There were 50 fish in each tank, and at the end of the experiment, a sample of 10 fish was

removed from each tank and their sperm cells examined for deformities. The data were the % of deformed sperm cells in each fish.

A t-test was then performed to test the H_0 that the addition of Pb in the sea water did not affect the % of deformed sperm cells in fish. The result showed a t value of 11.48, with 9 df and a P value of < 0.001.

Treatment Mean	% Deformities
Added Pb	35.7
Control	2.5

Q5-1. Discuss the design of this experiment, focusing on;

a. Adequacy of controls

b. Level of replication

Q5-2. Ignoring the above inadequacies, what would be your conclusions based on the results presented?

Question 6 - Two factor ANOVA

A biologist studying starlings wanted to know whether the mean mass of starlings differed according to different roosting situations. She was also interested in whether the mean mass of starlings altered over winter (Northern hemisphere) and whether the patterns amongst roosting situations were consistent throughout winter, therefore starlings were captured at the start (November) and end of winter (January). Ten starlings were captured from each roosting situation in each season, so in total, 80 birds were captured and weighed.

Format of starling.csv data files

SITUATION	MONTH	MASS	GROUP
S1	November	78	S1Nov
..	..	..	..
S2	November	78	S2Nov
..	..	..	..
S3	November	79	S3Nov
..	..	..	..
S4	November	77	S4Nov
..	..	..	..
S1	January	85	S1Jan
..	..	..	..

SITUATION Categorical listing of roosting situations

MONTH	Categorical listing of the month of sampling.
MASS	Mass (g) of starlings.
GROUP	Categorical listing of situation/month combinations - used for checking ANOVA assumptions

Open the starling data file.

Q6-1. List the 3 null hypothesis being tested

a. _____

b. _____

c. _____

Q6-2. Test the assumptions by producing **boxplots** and **mean vs variance plot**. **Note, use the variable GROUPS (which is a combination of SITUATION and MONTH for the assumption testing**

a. Is there any evidence that one or more of the assumptions are likely to be violated? (Y or N) _____ ☐

Q6-3. Now **fit a two-factor ANOVA model** and **examine the residuals**.

a. Any evidence of skewness or unequal variances? Any outliers? Any evidence of violations? ('Y' or 'N') _____ ☐ .

b. **Examine the ANOVA table** and fill in the following table:

Source of Variation	SS	df	MS	F-ratio	Pvalue
SITUATION	☐	☐	☐	☐	☐
MONTH					

	☐	☐	☐	☐	☐
SITUATION : MONTH	☐	☐	☐	☐	☐
Residual (within groups)	☐	☐	☐		

c. Copy and paste the ANOVA table from the Rcmdr output window into Word, format the table correctly and add an appropriate table caption.

Q6-5. An **interaction plot (plot of means)** is useful for summarizing multi-way ANOVA models. Summarize the trends using a **plot of means**.

Q6-6. In the absence of an interaction, we can examine the effects of each of the main effects in isolation. It is not necessary to examine the effect of MONTH any further, as there were only two groups. However, if we wished to know which roosting situations were significantly different to one another, we need to perform additional **multiple comparisons**. Since we don't know anything about the roosting situations, no one comparison is any more or less meaningful than any other comparisons. Therefore, a Tukey's test is most appropriate. Perform a **Tukey's test** and summarize indicate which of the following comparisons were significant (put * in the box to indicate $P < 0.05$, ** to indicate $P < 0.001$, and NS to indicate not-significant).

Situation 1 vs Situation 2 _____ ☐

Situation 1 vs Situation 3 _____ ☐

Situation 1 vs Situation 4 _____ ☐

Situation 2 vs Situation 3 _____ ☐

Situation 2 vs Situation 4 _____ ☐

Situation 3 vs Situation 4 _____ ☐

Q6-7. Generate a **bargraph** to summarize the findings of the ANOVA. **Save the graph (as a jpeg)** and **import the picture into Word**.

Q6-8. Summarize your conclusions from the analysis.

Question 7 - Two factor ANOVA

Here is a modified example from Quinn and Keough (2002). Stehman and Meredith (1995) present data from an experiment that was set up to test the hypothesis that healthy spruce seedlings break bud sooner than diseased spruce seedlings. There were 2 factors: pH (3 levels: 3, 5.5, 7) and HEALTH (2 levels: healthy, diseased). The dependent variable was the average (from 5 buds) bud emergence rating (BRATING) on each seedling. The sample size varied for each combination of pH and health, ranging from 7 to 23 seedlings. With two factors, this experiment should be analyzed with a 2 factor (2 x 3) ANOVA.

Format of stehman.csv data files			
PH	HEALTH	GROUP	BRATING
3	D	D3	0.0
..	..	..	..
3	H	H3	0.8

..	..	..	..
5.5	D	D5.5	0.0
..	..	..	..
5.5	H	H5.5	0.0
..	..	..	..
7	D	D7	0.2
..	..	..	..

PH   Categorical listing of pH (not however that the levels are numbers and thus by default the variable is treated as a numeric variable rather than a factor - we need to correct for this)

HEALTH   Categorical listing of the health status of the seedlings, D = diseased, H = healthy

GROUP   Categorical listing of pH/health combinations - used for checking ANOVA assumptions

BRATING   Average bud emergence rating per seedling

Open the stehman data file.

The variable PH contains a list of pH values and is supposed to represent a factorial variable. However, because the contents of this variable are numbers, R initially treats them as numbers, and therefore considers the variable to be numeric rather than categorical. In order to force R to treat this variable as a factor (categorical) it is necessary to first **convert this numeric variable into a factor**.

Q7-1. Test the assumptions by producing **boxplots** and **mean vs variance plot**. **Note, use the variable GROUPS (which is a combination of SITUATION and MONTH for the assumption testing**

     a. Is there any evidence that one or more of the assumptions are likely to be violated? (Y or N) ☐

Q7-2. Now **fit a two-factor ANOVA model** and **examine the residuals**.

     a. Any evidence of skewness or unequal variances? Any outliers? Any evidence of violations? ('Y' or 'N') ☐

     b. **Examine the ANOVA table** and fill in the following table:

Source of Variation	SS	df	MS	F-ratio	Pvalue
PH	☐	☐	☐	☐	☐
HEALTH	☐	☐	☐	☐	☐
PH : HEALTH	☐	☐	☐	☐	☐
Residual (within groups)	☐	☐	☐		

     c. **Copy and paste the ANOVA table from the Rcmdr output window into Word**, format the table correctly and add an appropriate table caption.

Q7-3. Summarize these trends using a **plot of means**.

Q7-4. In the absence of an interaction, we can examine the effects of each of the main effects in isolation. It is not necessary to examine the effect of HEALTH any further, as there were only two groups. However, if we wished to know which pH levels were significantly different to one another, we need to perform additional **multiple comparisons**. Since no one comparison is any more or less meaningful than any other comparisons, a Tukey's test is most appropriate. Perform a **Tukey's test** and summarize indicate which of the following comparisons were significant (put * in the box to indicate $P < 0.05$, ** to indicate $P < 0.001$, and NS to indicate not-significant).

pH 3 vs pH 5.5 ☐

pH 3 vs pH 7 ☐

pH 5.5 vs pH 7 ☐

Q7-5. **bargraph** to summarize the findings of the ANOVA. **Save the graph (as a jpeg)** and **import the picture into Word**.

Q7-6. Summarize your conclusions from the analysis.

Q7-7. Why aren't the 5 buds from each tree true replicates? Given this, why bother observing 5 buds, why not just use one?

Question 8 - Two factor ANOVA

An ecologist studying a rocky shore at Phillip Island, in southeastern Australia, was interested in how clumps of intertidal mussels are maintained. In particular, he wanted to know how densities of adult mussels affected recruitment of young individuals from the plankton. As with most marine invertebrates, recruitment is highly patchy in time, so he expected to find seasonal variation, and the interaction between season and density - whether effects of adult mussel density vary across seasons - was the aspect of most interest.

The data were collected from four seasons, and with two densities of adult mussels. The experiment consisted of clumps of adult mussels attached to the rocks. These clumps were then brought back to the laboratory, and the number of baby mussels recorded. There were 3-6 replicate clumps for each density and season combination.

Format of quinn.csv data files

SEASON	DENSITY	RECRUITS	SQRTRECRUITS	GROUP
Spring	Low	15	3.87	SpringLow
..	..	..	..	..
Spring	High	11	3.32	SpringHigh
..	..	..	..	..
Summer	Low	21	4.58	SummerLow
..	..	..	..	..
Summer	High	34	5.83	SummerHigh
..	..	..	..	..
Autumn	Low	14	3.74	AutumnLow
..	..	..	..	..

SEASON	Categorical listing of Season in which mussel clumps were collected independent variable

DENSITY	Categorical listing of the density of mussels within mussel clump independent variable
RECRUITS	The number of mussel recruits response variable
SQRTRECRUITS	Square root transformation of RECRUITS - needed to meet the test assumptions
GROUPS	Categorical listing of Season/Density combinations - used for checking ANOVA assumptions

Open the quinn data file.

Confirm the need for a square root transformation, by examining **boxplots** and **mean vs variance plots** for both raw and transformed data. Note that square root transformation was selected because the data were counts (count data often includes values of zero - cannot compute log of zero).

Q8-1. Now **fit a two-factor ANOVA model** (using the square-root transformed data and **examine the residuals**.

   a. Any evidence of skewness or unequal variances? Any outliers? Any evidence of violations? ('Y' or 'N') _____ ☐.

   b. **Examine the ANOVA table** and fill in the following table:

Source of Variation	SS	df	MS	F-ratio	Pvalue
SEASON	☐	☐	☐	☐	☐
DENSITY	☐	☐	☐	☐	☐
SEASON : DENSITY	☐	☐	☐	☐	☐
Residual (within groups)	☐	☐	☐		

Q8-2. Summarize these trends using a **plot of means**. Note that graphs do not place the restrictive assumptions on data sets that formal analyses do (since graphs are not statistical analyses). Therefore, it data transformations were used for the purpose of meeting test assumptions, it is usually better to display raw data (non transformed) in graphical presentations. This way readers can easily interpret actual values in a scale that they are more familiar with.

Q8-3. The presence of a significant interaction means that we cannot make general statements about the effect of one factor (such as density) in isolation of the other factor (e.g. season). Whether there is an effect of density depends on which season you are considering (and vice versa). One way to clarify an interaction is to analyze subsets of the data. For example, you could examine the effect of density separately at each season (using four, single factor ANOVA's), or analyze the effect of season separately (using two, single factor ANOVAs - known as simple main effects ANOVAs) at each mussel

density.

For the current data set, the effect of density is of greatest interest, and thus the former option is the most interesting. Perform the **simple main effects ANOVAs**.

   a. Effect of density in Autumn: (F = ☐, df = ☐, ☐, P = ☐)

   b. Effect of density in Spring: (F = ☐, df = ☐, ☐, P = ☐)

   c. Effect of density in Summer: (F = ☐, df = ☐, ☐, P = ☐)

   d. Effect of density in Winter: (F = ☐, df = ☐, ☐, P = ☐)

Q8-4. What conclusions would you draw from these findings?

   a. Was the effect of DENSITY on recruitment consistent across all levels of SEASON? (Y or N) ☐

   b. How would you interpret these results?

Welcome to the end of Worksheet 4