BIOLOGY 1001 LAB 7 BIOLOGICAL DATA II

INSTRUCTIONS:

1. During this lab it is important to have RStudio (or RStudio Cloud), the "Quantitative skills for biology" guide and the Lab 7 Worksheet all open on your computer at the same time.

2. The activities/questions in the worksheet come from 2 Quantitative skills laboratory (https://ahurford.github.io/BIOL-1001/lab2.html). We are presenting them in this worksheet to lessen confusion and for ease of grading.

3. DO NOT copy and paste information directly from 2 Quantitative skills laboratory , "Quantitative skills for biology" or other online materials when answering your questions. **This will be considered plagiarism and you will receive a grade of zero with possible further disciplinary action as per the University Calendar.** Your answers must be in your own words. (**This, of course, does not include the cutting and pasting of your code and graphs from your R script or console.)**

4. SAVE your document frequently. Once completed, save your assignment as a PDF document and upload to the Assignment Folder entitled "Lab 7 Biological Data II" in the pull-down menu under Assessments.

5. If you have any questions, please contact your lab instructor.

6. If you experience technical difficulties with your upload, contact the Client Support Team of the Centre for Innovation in Teaching and Learning at www.citl.mun.ca/support

Notice from Memorial University:

You agree to the following; "All members of the Memorial University of Newfoundland community, including students, faculty, and staff, shall treat others with respect and fairness, be responsible and honest, and uphold the highest standards of academic integrity.

By submitting this assignment, I unequivocally state that all work is entirely my own and does not violate Memorial University's Academic Integrity policy."

EXERCISE 1. Data entry and graphing with a continuous independent variable

1. **Enter the data**
   Include a copy of the code you used in your lab report. (***Copy and paste the code you used to import this dataset below.) (1.0 mark)***

From 8.2.1 (code or import function)

Code:

```
LangdonDataset <- read.csv("LangddonDataset.csv")
```

Import function:

```
LangdonDataset <- read.csv("~/Library/Mobile Documents/
com~apple~CloudDocs/Documents/career/LI/_B1001 F22/B1001 lab 7 R2/
LangdonDataset.csv")
```

2. **Make a graph**
   Follow the instructions in Making graphs in R to make a scatterplot for these data and replicate the figure from the Campbell textbook (see Figure 14.1 below). Include both the code you wrote and the final figure in your lab report. NOTE: Don't worry about figuring out how to write superscripts and subscripts in R for the axis-labels. **You can use the "^" symbol to indicate a superscript, and the "-" to indicate a subscript**.
   Copy and paste both the code you wrote and the final figure below. (3.5 marks, 2.5 for code and 1 for graph)

   Read and follow "NEW TIP" before answering.

As per the "NEW TIP", easiest method is to create a vector from the data frame:

```
CO3<-LangdonDataset$CO3
G<-LangdonDataset$G
```

Then use the plot function from last week, but fix vector names and x and y labels

plot(temp, fishA, pch = 0, xlab = "temperature (degrees C)", ylab = "opercular beats (bpm)")

And the **CODE** is:

```
plot(CO3, G, pch = 0, xlab = "[CO3^2] (umol/kg of seaweed)",
ylab = "Calcification rate (mmol CaCO3/m2 per day)")
```

***NOTE*: In the lab it makes adding the regression line sound optional (it's worded as "if you want to…"). Don't penalize if they don't include abline, summary and coefficients. A lot of them won't have had stats yet. The code is below:**

```
abline(lm(LangdonDataset$G ~ LangdonDataset$CO3))
summary(lm(LangdonDataset$G ~ LangdonDataset$CO3))
```

And the results are:

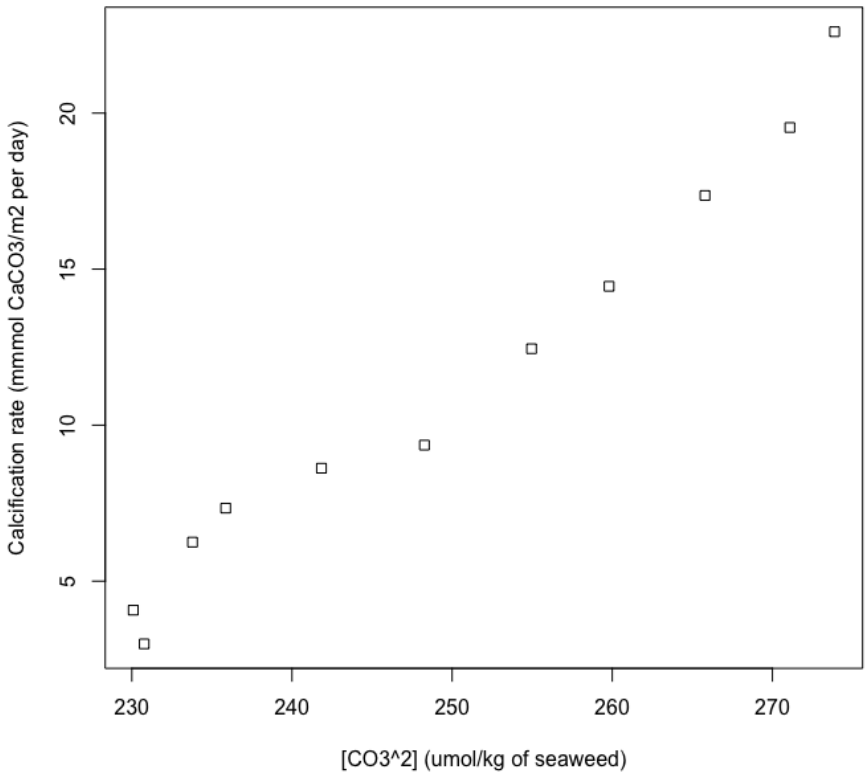
```
Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)      -85.73363    5.44378  -15.75 7.38e-08 ***
langdon_data$CO3   0.38895    0.02176   17.87 2.44e-08 ***

Adjusted R-squared:  0.9695
```

***NOTE:* This code (the command expression(paste()) is the way to make the Greek "mu" symbol appear, but we didn't teach them that and it would be fairly tricky to figure out on their own. Their caption text may be slightly different but as long as it's reasonably descriptive and has units, give full marks. Deduct 0.5 mark for missing the units.**

***NOTE:* If they construct their vectors by simply listing the data points rather than using $ deduct 1 mark.**

And the GRAPH is:



3. **Make a graph**

Follow the instructions in Making graphs in R to make a second graph - a line graph for these data. Export your line graph insert it into your lab report. Write a figure caption. Include both the code you wrote and the final figure in your lab report.

Export your line graph and paste it below. Write a figure caption for your graph. Also copy and paste the code you wrote to create your line graph. (4.0 marks, 2 for code, 1 for graph, 1 for figure legend)

HINT: Your figure caption can be typed into this document below your figure

Same code as above, just add the line function (type="l")

plot(CO3, G, pch = 0, xlab = "[CO3^2] (umol/kg of seaweed)", ylab = "Calcification rate (mmol CaCO3/m2 per day)", type="l")

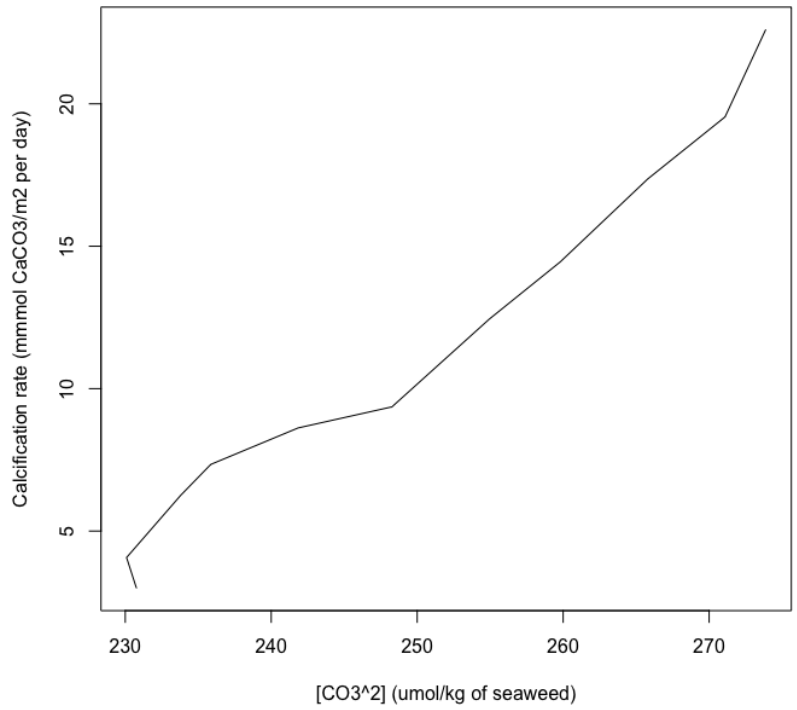


Figure 2. Relationship between net community calcification rate and concentration of carbon ions (CO_3^{2-}) during April-June, 1998 at the BIOSPHERE-2 Center. The net calcification rate increases linearly as the concentration of carbon ions increases.

***NOTE*: For the title of the figure, students may not be as detailed as above. It can be as simpleas "The effect of the concentration of carbon ions on the net community calcification rate calcification rate". Again, do no penalize students if they do not use the symbol µ. They may simply write "mu"**

EXERCISE 2. Bar plots and boxplots with discrete independent variables

1. Follow the instructions in Entering and loading data to load the data file FalikDataset.csv using the command line (read.csv) option. **Note** that the data in the CSV file are ordered differently than in the list above (to verify this you can open the CSV file in Excel to see the full set). **Copy and paste the code you used to import this dataset below. (1.0 mark)**

From Exercise 1 above (code or import function)

Code:

```
FalikDataset <- read.csv("FalikDataset.csv")
```

Import function:

```
FalikDataset <- read.csv("~/Library/Mobile Documents/com~apple~CloudDocs/
Documents/career/LI/_B1001 F22/B1001 lab 7 R2/FalikDataset.csv")
```

2. Looking at the dataset, can you tell how many treatments were applied in the experiment? How many plants were used for each treatment? **(1.0 mark)**

# of Treatments: 2 (control and 15 minutes)
# of plants: 11 for each

3. What is the smallest stomatal width measured? What is the largest? **(1.0 mark)**

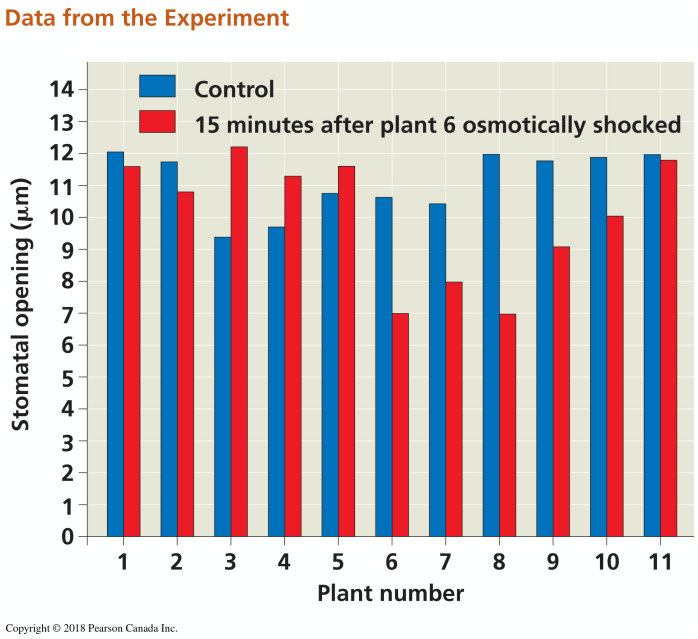
Smallest Stomatal Width: 7.081 µm
Largest Stomatal Width: 12.110 µm

EXERCISE 2b. Making a bar plot

1. Follow the instructions below to replicate the figure from the Campbell textbook (barplot; see Figure 14.3 below).

Copy and paste your code and bar plot below. (5 marks; *3.0 marks for the code, 1.0 mark for the graph, 1.0 mark for figure title*)

Don't forget to include your figure caption. It can simply be typed into this document after you have pasted your figure.



GIVEN:

```
width <- falik_data$width #This creates a vector of data of the plant stomatal widths
plant <- falik_data$plant #The creates a vector of data of the plant ID numbers

barplot(width, names = plant, col = c("blue", "red"))
```

CODE:

```
barplot(width, names = plant, col = c("blue", "red"),
    xlab = "Plant #",
    ylab = "Stomatal opening (um)",
    legend = TRUE, ylim = c(0,20))

legend("topleft", fill =c("blue", "red"),
   legend=c("Control", "15 minutes after osmotic shock"),
   title="Treatment", bty = "n")
```

***NOTE*: They were NOT taught how to do a legend so don't deduct marks if these lines of code are not there. As long as they describe what the two colours mean in the figure caption, they should get full marks. If legend or sufficient description is missing, deduct 0.5 mark.**

GRAPH:

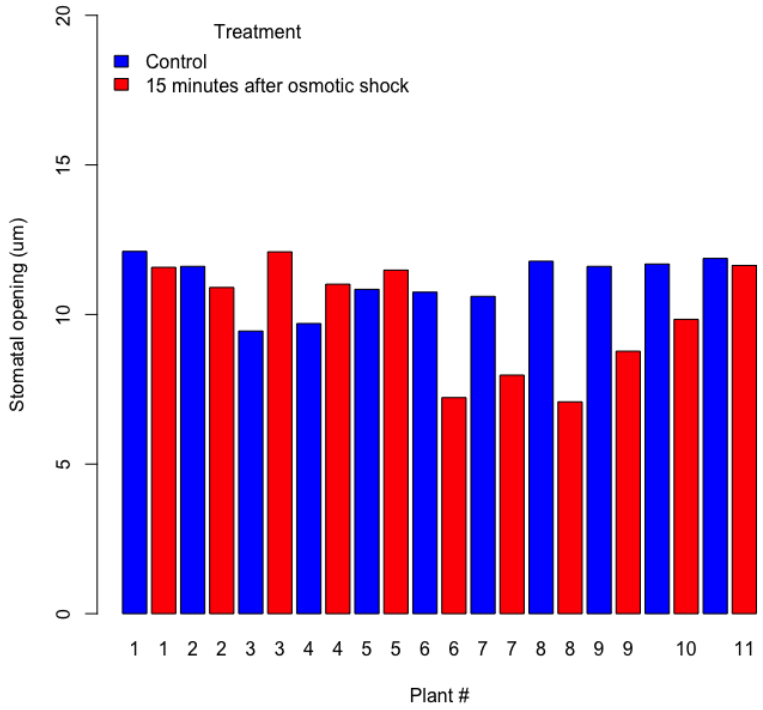


Figure 3. Stomatal width of *Pisum sativum* plants in response to osmotic stress. The stomatal response was measured 15 minutes after drought was induced to plant number 6. Blue bars represent plants in the control treatment (unshared roots) and red bars show plants that shared roots with the stressed neighbour or with a unstressed neighbour (shared roots).

***NOTE*: For the title of the figure, students may not be as detailed as above. It can be as simple as "The effect of the osmotic stress on stomatal width of control and experimental *Pisum sativum* plants".**

2. In the bar plot, what can you tell about the stomatal openings of the treatment vs. the control plants? Is it consistent across all individual plants? (**1.0 mark**)

The plants in the treatment group show smaller values and greater variance of the mean stomatal opening when compared with plants in the control treatment. However, this is not consistent across all plants because some show less variation in the stomatal opening width.

3. Follow the instructions below to re-plot the data from EXERCISE 2b as a boxplot that compares control versus treatment.

Export your box plot and paste it below. Include a figure caption (2.0 marks)

NOTE: If you describe the meaning of the symbols/colours in the caption you do not need a legend in the box plot. If you want to try adding a legend (adding one is optional) then look up "legend" in the help window.

GIVEN:

```
control <- subset(falik_data, type == "control") #This subset command is a handy function to break the data into subsets. Here, we're creating a data frame with just the values from the control.
treatment <- subset(falik_data, type == "15min") #Here we are subsetting the data to contain only the data from the treatment

boxplot(control$width, treatment$width, names = c("control", "treatment"), ylab = expression(paste("stomatal opening ( ", mu, "m")))
```

CODE IS NOT REQUIRED (as it is given to them above):

GRAPH:

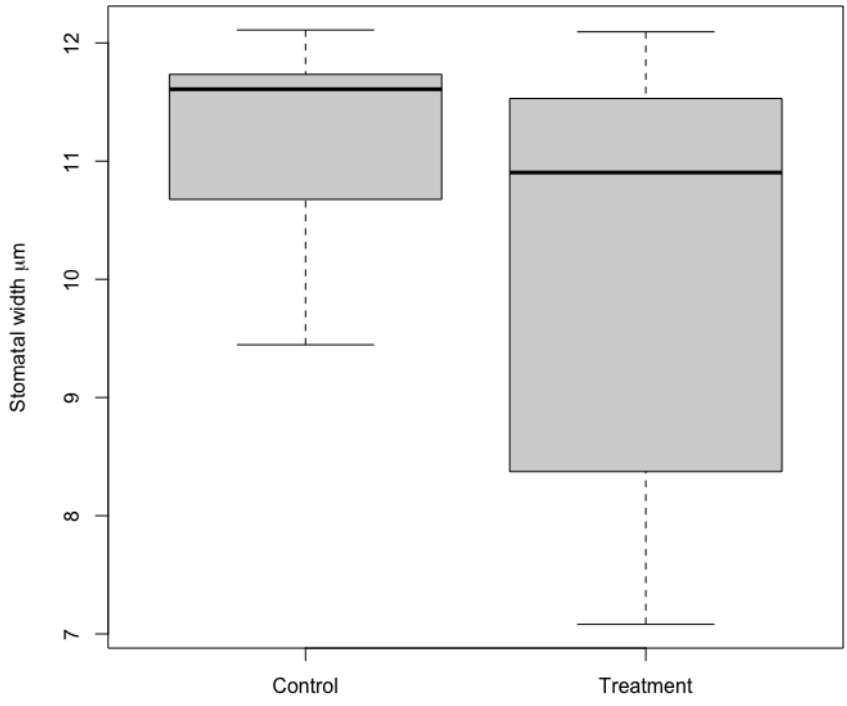
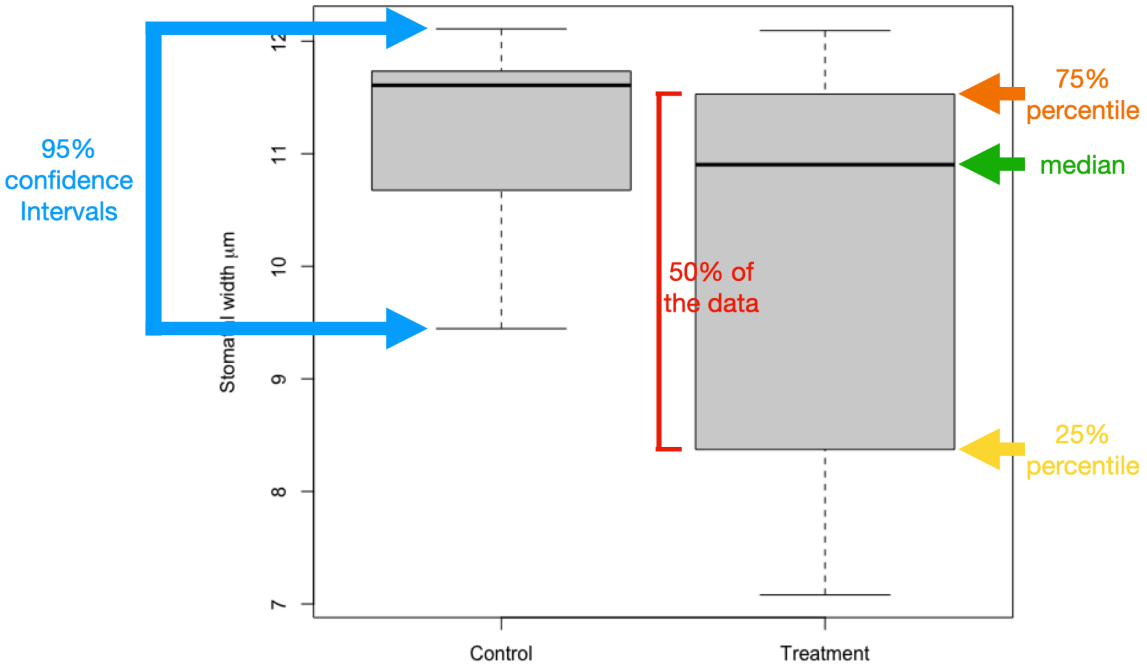


Figure 4. Stomatal width of *Pisum sativum* plants in response to osmotic stress. The stomatal response was measured 15 minutes after drought was induced. Treatment plants (with shared roots) show a smaller median and larger variance relative to control plants (with unshared roots).

***NOTE*: For the title of the figure, students may not be as detailed as above. It can be as simple as "The effect of the osmotic stress on stomatal width of control and experimental *Pisum sativum* plants".**

4. Label the following on the boxplot: **median**, **25%** and **75% quantiles**, **95% confidence limits**. **(2.5 marks, 0.5 per label and 0.5 for labelling both plots)**

   HINT you can insert the exported graph into a blank PowerPoint slide and use lines/text boxes to draw the labels. Then save the slide as an image file (e.g., JPEG) and insert it in your lab report.



5. What does the boxplot tell you about the differences between the treatment plants (plants 6-11) vs. the control plants that the bar plot does not? **(1.0 mark)**

   In the boxplot we can see how the distribution of the stomatal width differs between the treatment group and the control group. ***(Could also mention that, the median stomatal opening of the treatment group is lower than the median for the control group, or that the variance of the treatment plants is larger relative to the control group.)***

6. Why do you think a scientist might choose to represent these data one way over another? **(2.0 marks)**

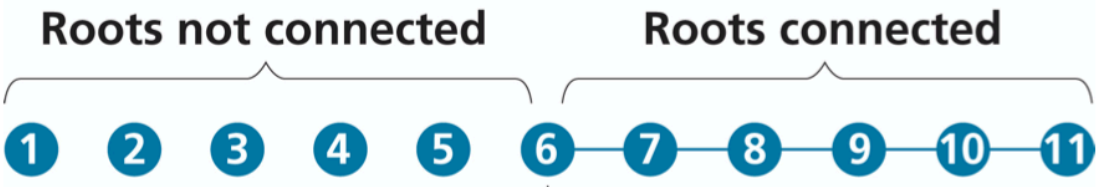
The bar plot summarizes the data by means and standard deviations or standard errors (e.g. the mean of all the replicates of plants on position 3). A scientist may choose to represent data with a bar plot if the underlying distribution of the data is the same for all the groups compared, and if the individual data points are not relevant.

In most other cases a boxplot would be a better representation because it shows the distribution of all the observations according to group. Therefore, a scientist may choose to use a boxplot to emphasize how the median, the interquartile range and the outliers compare among treatment groups within the dataset.

NOTES:
Bar plot would be a good idea if the researcher is interested in looking at how each of the 11 individual plants behaved after 15 minutes and comparing one plant to the rest of the others in the experiment.

The box plot groups the data into two distinct groups: control and treatment. It allows a comparison between the groups as opposed to individuals. IT also conveys much more information such as the medians of each group and the variance/variation within each group.

Control (no osmotic shock)



Treatment (15 min of osmotic shock)