Homework 1

Formatting: There are two formatting options for your assignments:

1. A single document (MS Word or equivalent) that alternates between your annotated R code and the corresponding requested output (figures, model output, etc). I.e., format it like the lecture notes.
2. An annotated R script that reproduces your results, and a single document (MS Word or equivalent) that includes all of the requested output (figures, model output, etc).

The file "ManyakBellSotka…" contains the raw data from the attached 2013 Am Nat paper. This study looked at the causes underlying Bergmann's rule (within a species, animals tend to have a larger body size at higher latitudes) in an isopod. Here you'll use their data to recreate some analyses looking at how body size and life history characteristic vary between northern and southern populations.

1. (3 pts) From the "Lab Cultures" worksheet, plot the mean length (Length..mm.) by Sex and Region. Plot error bars for +/- 1 standard error of the mean. There are many ways to do this; some options for summarizing data in a dataframe include using tapply(), or using ddply() and summarize() from the "plyr" package. To plot the means and error bars you can combine plot() and segments(), or google one of many functions people have written to do it.

2. (4 pts) Use lm(), and F-tests on the model, to test whether there is a significant difference between Sexes in mean length, whether there is a significant difference between Regions in mean length, and whether the effect of Region differs between Sexes. Report F-statistics and degrees of freedom. How do you interpret these results?

3. (3 pts) Plot the model-fitted group means and standard errors (using the "effects" package is probably easiest).

4. (3 pts) Using the "Egg Data" worksheet, plot Number of Eggs vs. Length, and color code the plotted points by Region. (Note, the authors did not include a Region variable in this worksheet, so you will need to create one by referring to the other worksheets). There are many ways to do this; this thread has some suggestions http://stackoverflow.com/questions/7721262/colouring-plot-by-factor-in-r

5. (5 pts) Fit a linear model to test whether the relationship between Number of Eggs and Length differs between Regions. Perform and report the appropriate F-test to test this question. How might you interpret these results?

6. (3 pts) For #5 you have essentially fit two linear regressions, one for each Region. Take your plot from #4 and add these two lines. abline() will help.

7. (4 pts) Model diagnostics: for the model you fit in #5, make some plots that explore whether the residuals of the model are normally distributed, whether the variance of the residuals increases as Length increases, and whether the variance of the residuals varies between regions.