Assignment 2: Linear Models

Due: **Friday, March 11, 2022**

Obtain a data-set and analyze it by fitting a linear model in R (mixed effects model, if necessary and generalized linear models are also ok, but we won't cover these in class until March 8).

- Obtain a data set (from your supervisor or online data depository such as Dryad or your own if you have not analyzed it yet). Include just one response variable.

- For the explanatory variables, include at least one proper fixed effect, such as an experimental or observational treatment. Can be categorical or numeric.

- Include at least 1 but not more than 2 additional explanatory variables (random or fixed factors, blocks, covariates, etc).

- Explain (in a paragraph) the purpose of the study that yielded the data.

- Explain the specific data set you are using. For example, say where the data are from, give the meaning of the variables, and so on.

- Show the raw data in a figure.

- State what hypotheses you will test with these data.

- State what parameters you will estimate with these data.

- Fit a linear model to the data in R. Explain in words the model you fit.

- Interpret the output. To assess biological significance, explain the parameter estimates. What do they mean? To assess statistical significance, explain the null hypotheses and interpret any relevant test results.

- Show the model graphically (ideally, with the data). Explain what this/these graphs is/are showing. This can be similar to your initial graph of the data, but with the model fit added.

- Address how well the statistical assumptions of your analysis were met, and how you handled violations.

- State the overall conclusions reached from your analyses of biological and statistical significance.

- Try to keep sample size manageable to avoid computational issues (use subset if real data set too huge).

- Include your clean, annotated R code in an appendix.

- Include your dataset in an appendix.

Email your assignment to Will Stein (rwstein@gmail.com) and myself (lmgonigl@sfu.ca) as a single pdf file with filename:
LASTNAME_FIRSTNAME_ASSIGNMENT2.pdf

Rubric

1. Modeling [**5 points**]

   - Appropriateness of the model in relation to structure of data and goals of analysis
   - Analysis of assumptions
   - Correct use of coefficients, etc

2. Interpretation of results [**5 points**]

   - Clear interpretation of findings, estimates, etc
   - Quality of conclusions

3. Visuals [**5 points**]

   - Understandable dataset
   - Quality of graphs showing model results and model fit (e.g., residual plots)

4. Quality of your R code [**5 points**]

   - Annotated
   - Readable
   - General organization