Analysis of Ecosystems homework week 6

Multiple regression & ANOVA

The assignment

Data preparation

In this assignment you will fit two models:

- A multiple linear regression model with a continuous response variable, a continuous predictor variable, and a categorical predictor variable.
- An analysis of variance (ANOVA) model with a *different* categorical predictor variable.

Identification

Use a dataset of your choice. You can use `mpg` data from `ggplot2` if you want, any other from `data()`, or data of your own. *Just make sure you aren't using any of the same variables from the same data we've used in class.*

In Code Chunk 3, show the structure of the dataset you're using and identify which variables you use. Below the code chunk, describe what you know about the data and the variables you've selected.

Assumptions

Starting with Code Chunk 4 and copy-pasting as necessary, present visual and/or statistical tests of model assumptions relevant *before* models are fit.

Multiple linear regression

Graphing

Using `ggplot` in Code Chunk 5, make an appropriate graph containing no less than three variables. Use color and shapes that present data in a clear way legible regardless of whether or not one is able to use color to distinguish variables.

Hypothesis statement

As explicitly as possible, state:

- A null hypothesis
- An alternative hypothesis that refers to two variables simultaneously.

Fit a linear model

Use Code Chunk 6 to fit a multiple regression model and present a summary of the results.

Interpret

Summarize the results and present a conclusion with specific reference to the hypotheses you stated above (if you get here and didn't state hypotheses as described above, better go back and do that.)

ANOVA on categorical predictor variables

Graphing

Using `ggplot` in Code Chunk 7, make an appropriate graph comparing the continous variable against the new categorical variable.

Hypothesis statement

As explicitly as possible, state:

- A null hypothesis
- An alternative hypothesis

Fit an ANOVA model

Use Code Chunk 8 to fit an ANOVA model and present a summary of the results. The statistical analysis must differentiate among specific groups, so be sure to check the number of groups and ensure that you present sufficient analysis to distinguish pairwise contrasts.

Interpret

Summarize the results and present a conclusion with specific reference to the hypotheses you stated above.