Class Exercise 3: Fantastic Beasts
Physiology & Life History Traits for All Living Mammals
Due April 5, 2018 by 5pm via Sakai

SUBMISSION GUIDELINES

Please name your report: BISC111_ClassEx3_YOURBANNERID.docx/.pdf, and include your Banner ID on the first line of your document. Do not include your name in your document, because I prefer to grade anonymously. Your Data Exercise is due by 5pm on Thursday, April 5th.

INTRODUCTION

In this Data Exercise, you will use an app built by Prof. Matthes to visualize relationships between variables within a large dataset (called PanTHERIA) of metabolic rates, body mass, and other life history traits for all living and recently extinct mammals. The dataset includes information about 100,740 species of mammals, and is organized by mammal order (of taxonomy) and is described in detail within the journal article Jones, *et. al.* (2009) (also linked within the app). You will use the app to visualize and explain relationships among variables describing the adaptive traits of mammals and investigate outlier taxa that defy relationships otherwise well described by the data. The data app is web-based at: http://r.wellesley.edu:3838/apps/bisc111/mammal-app/
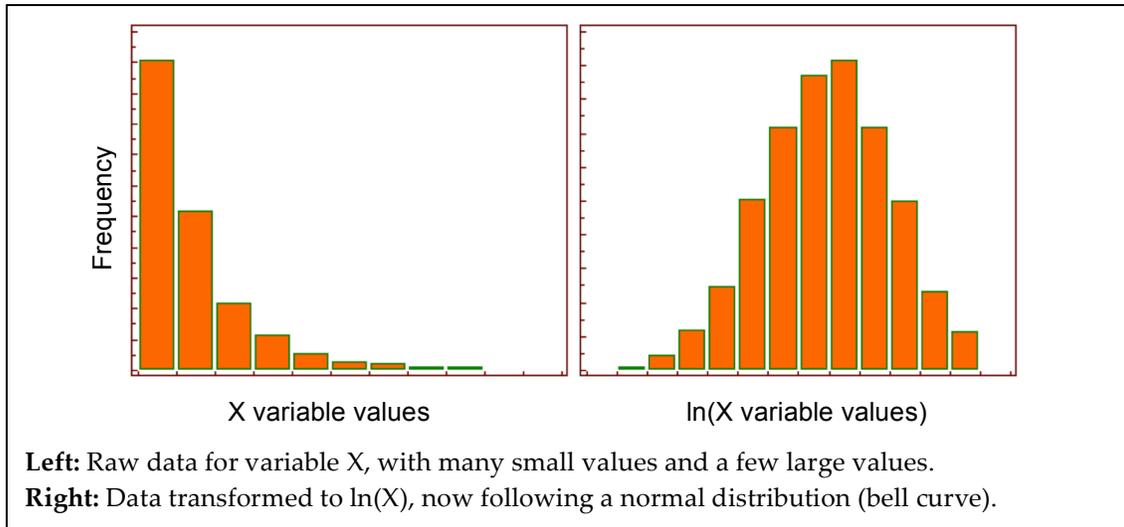
You use the app by choosing x- and y-variables to plot from the PanTHERIA dataset. There is an option to plot data across all orders of mammals or within particular orders (primates, rodents, etc.) to examine patterns within a smaller taxonomic group. Some orders have relatively little data, so if you get an error while trying to create a graph data for a particular order, it's likely because there aren't enough data to create a plot and you should choose a different order or all orders.

USING LOGARITHMS TO DISPLAY AND MODEL DATA

All of the data in this app are plotted on a logarithmic scale. In contrast to linear scales, logarithmic scales are particularly helpful for visualizing data with values that bridge many orders of magnitude[1]. If we were to display the data in this app on a linear scale, the relationship between variables would look very non-linear (exponential). A logarithmic scale better fits most animal traits than a linear scale (especially when we

---

[1] An order of magnitude is adding/removing zeros to a number, so 10 is an order of magnitude bigger than 1, and 0.1 is an order of magnitude smaller than 1.

consider all mammals) because animals bridge such a huge range in the magnitude of trait values (for example, size).

Left: Raw data for variable X, with many small values and a few large values.
Right: Data transformed to ln(X), now following a normal distribution (bell curve).

Using logarithms isn't just useful for visualizing mammal trait data, it's also useful for modeling the relationship between traits. This is especially true for linear regression (fitting a straight line to the data) since linear regression assumes that the error between the data and the model fit (residuals) are normally distributed (follow a bell curve). In this app, you can choose to fit a linear model between the two log-transformed variables that you plot on the x- and y-axis. The general equation for a linear model (a straight line) is:

$$Y = mX + b$$

where m is the slope of the line, and b is the intercept of the line, and X are the values of the variable plotted on the x-axis, and Y are the values of the variable plotted on the y-axis. The equation for a line fit to the log-transformed data within this app is:

$$\ln(Y) = m \ln(X) + b$$

where all of the variables are the same as the equation above. You can see the fit of the linear model below the graph, where the slope and intercept are calculated for the X and Y variables displayed within the graph. This section also displays the R-squared value (R2 in the app) for the fit. The R-squared value describes the proportion of the variance in the dependent (Y) variable that is explained by the independent (X) variable. It ranges from 0 to 1, where 0 means that X doesn't explain any variability in Y, and 1 means that X perfectly predicts Y. You can also see the number of data points that were included for a particular model (this changes as you change the variables on the right).

EXPLORING DATA OUTLIERS

You can also click and drag a box around data points on the graph to pull up the species names that those data represent; this shows up in a table below the graph and below the linear model description (you might have to scroll down in the window to see this). You can copy and paste species names from the app and search for them on the Internet to get more information about what types of traits the highlighted animals have (Wikipedia is generally a good source for species information).

CLASS EXERCISE QUESTIONS (5 POINTS EACH)

1. Use the app to construct a graph with body mass on the x-axis and neonate (newborn offspring) mass on the y-axis with All Orders of mammals. Fit a linear model to the data. Include a screenshot of your final graph, and report the slope, intercept, R-squared value, and number of data points. Explain why you would expect to see a relationship between adult and neonate mass for mammals.

2. In the graph from Question 1, investigate more closely the green/yellow points that represent Orders that fall well below the linear model line for All Orders. Identify the orders that these data represent, and speculate as to why these taxa might be outliers from the rest of the data (you can do an internet search to learn what types of animals are included in these outlier data). Make a graph with just one of these outlier Orders, and compare the slope and intercept of the fit within this order to the slope and intercept from Question 1.

3. Look through the list of possible x- and y-variables that you can investigate. Choose two that you think might be connected as fitness trade-offs, and hypothesize a potential relationship that might exist between the two of them (across 'All' orders of mammals). In your answer, be clear about the mechanisms connected to fitness that you expect would control the pattern that you hypothesize.

4. Graph the relationship that you proposed in Question 3. Fit a linear model to the data. Include a screenshot of your final graph, and report the slope, intercept, R-squared value, and number of data points. Do the data support your hypothesis?

5. Investigate whether the relationship from Question 4 holds within individual orders (i.e., within Primates, within Carnivores, etc.). Pick two individual orders, graph the same relationship explained in Question 4, and explain whether these results are consistent or inconsistent with the relationship across 'All' orders. If the relationship for a particular order is inconsistent, speculate as to why this might be the case.