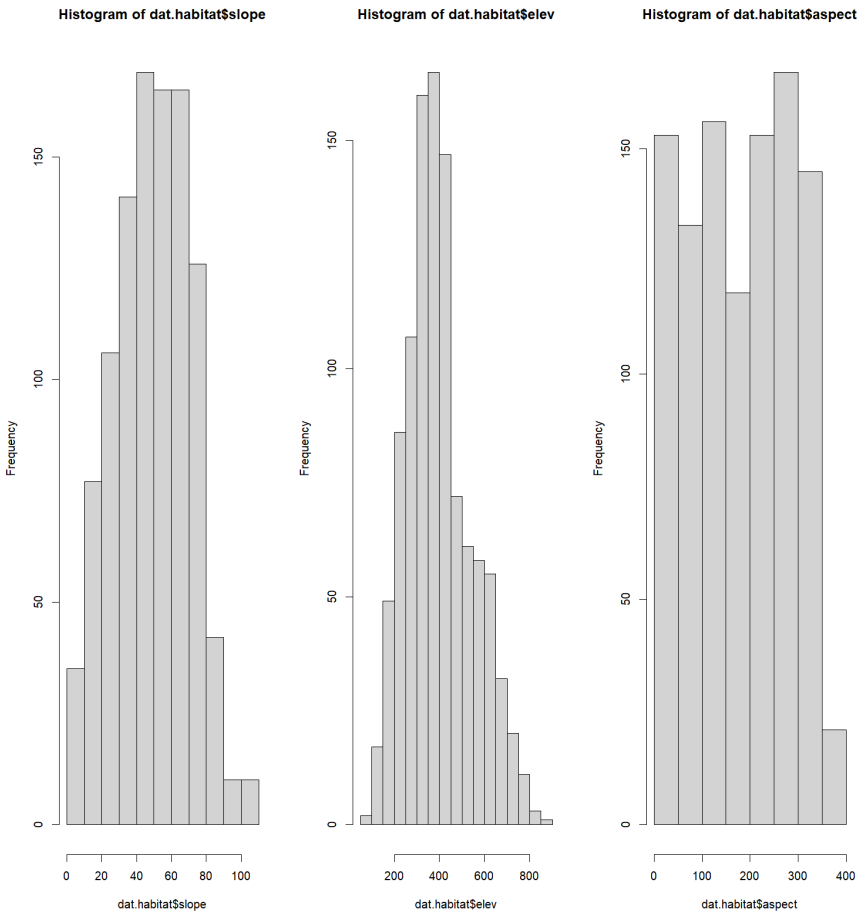
Matthew Jusino

Data Exploration & Deterministic Functions Assignment

Q1. Create histograms for the three terrain variables: elevation, slope, and aspect. Plot all three histograms in one figure and include it in your report.



Q2. Consider the distribution of elevations at the bird census sample sites.

- Interpret the shape of the elevation histogram in non-technical language that a non-scientist audience would understand. Some points to consider:
- Are there more high- or low-elevation sampling sites?
- Is there an even distribution of sampling site elevation?

Your answer should be 1-2 short paragraphs in length.

The elevation histogram shows the frequency of the occurrence of sites at particular intervals of elevation. It is skewed to the left, showing that there are significantly more sites at lower elevations than there are at higher elevations. You can see that the “middle” of the histogram is over about 400 m elevation, as that’s where the highest of the bars are clustered, and it drops off to either side. There is

not an even distribution. While the elevations of sites reach over 800 m, there are notably fewer of those than sites down around the 400 m range.

Q3. What are the units of slope in this data set?

The units of slope are in percentage (percent slope).

Q4. Consider the distribution of slopes at the bird census sample sites.

Interpret the shape of the slope histogram in non-technical language that a non-scientist audience would understand. Some points to consider:

Are most sample sites flat?

Is there an even mixture of steep and shallow slopes?

Your answer should be 1-2 short paragraphs in length.

The slope histogram shows clustering primarily around the 30-70% slope range. Most sample sites have a slope somewhere in this range. There are very few sites that are flat or nearly flat, with most having some sort of gradient. There is not an even mix of shallow and steep slopes, though it comes moderately close. There is a slight skew towards shallower slopes, as evidenced by the fact that outside of the three tallest bars, there tend to be more and taller bars to the left of the tallest bars, showing that there are more shallow gradient sites.

Q5. Briefly define aspect, describing the units used in this dataset.

Aspect is given in degrees, on the typical scale of 0-360. Aspect means the direction the land faces with respect to magnetic north, represented by 0/360.

Q6. Consider the distribution of aspect at the bird census sample sites.

Interpret the shape of the aspect histogram in non-technical language that a non-scientist audience would understand. Some points to consider:

Do the sampling sites tend to be on north-facing slopes?

South-facing?

Evenly distributed?

Your answer should be 1-2 short paragraphs in length.

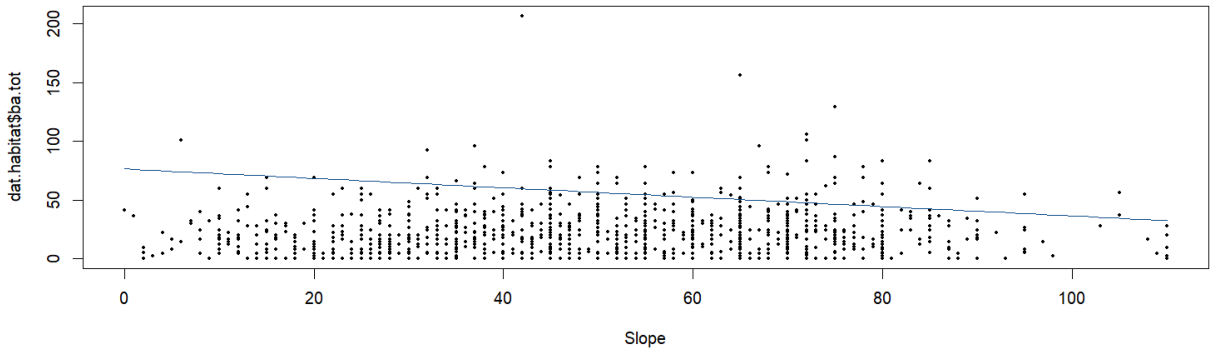
The aspect histogram shows a fairly even distribution of directions faced by the sites. There are slightly more north facing sites, as shown by the high frequency of sites close to both 0 degrees and 360 degrees. There are also slightly fewer sites that face south, as the bar close to 180 degrees is slightly shorter than the surrounding bars. Other than that, the sites are fairly evenly distributed around the compass directions.

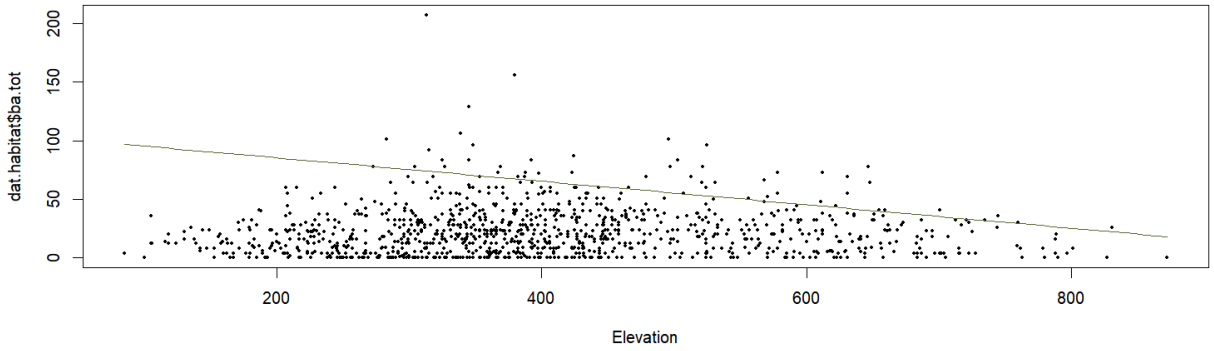
Q7. Create scatterplots of total basal area and each of the the terrain variables: elevation, slope, and aspect.

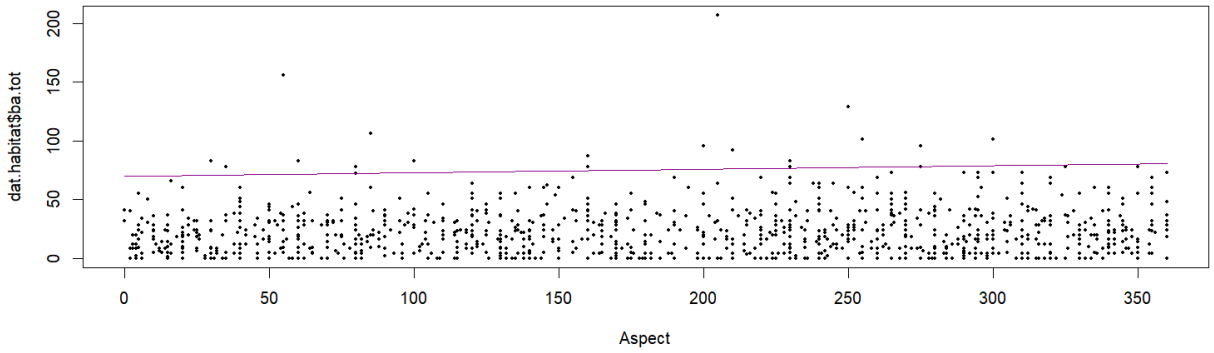
Basal area should be on the y-axis.

Visually inspect the plots and fit a linear function to each terrain variable.

Review the linear model parameterization section of the assignment walkthrough if needed.







Q8. For each terrain variable (elevation, slope, aspect), describe the relationship you observe and your model fit. You should consider

- Is there a noticeable *association*?
- If so, is it *linear*?
- Based on a visual assessment, is your linear model a good fit for the data, why or why not?

There appears to be a slight association between slope and basal area, but I don't think a linear model is a good fit for it, as it appear to be strongest around the slope of 40%, falling off to either side, making a linear model a bad fit as it cannot represent points on both ends. There appears to be a stronger association between elevation and basal area, with basal area decreasing as elevation increases. I think a linear model is still not the best fit, as there appears to be a peak in correlation around 300-400 m, so a linear model also leaves out points in this plot. There appears to be little to no association between aspect and basal area. This makes a linear model a good fit for the data, as it can easily describe the general lack of association between the two variables.