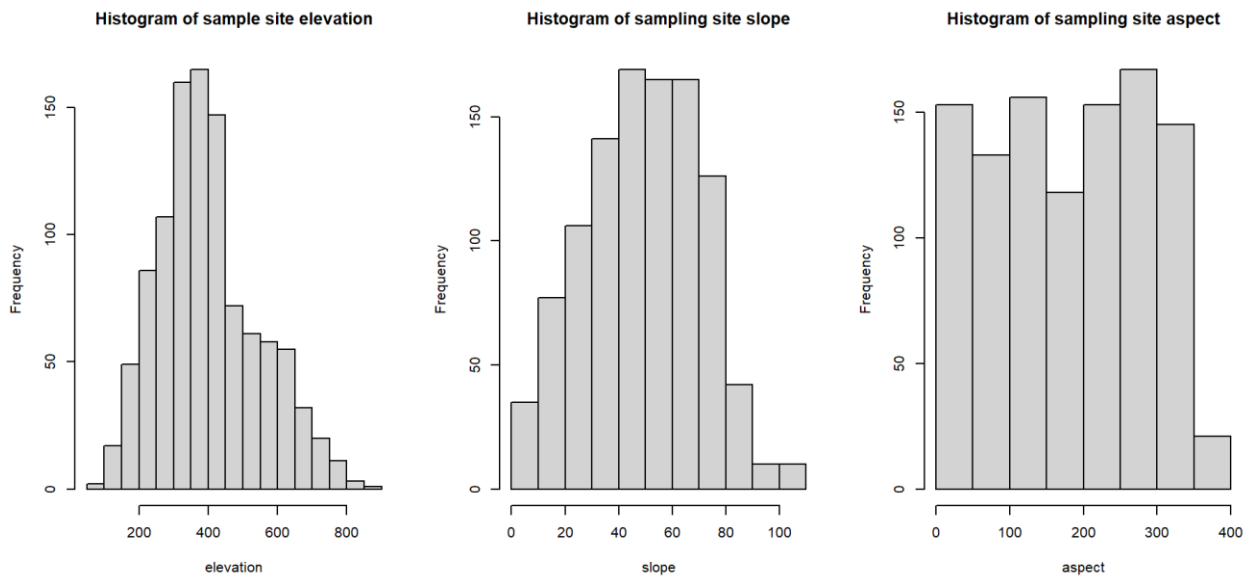


Julian Burgoff

9/28/2022

Analysis of Environmental Data

Data Exploration and Deterministic Functions

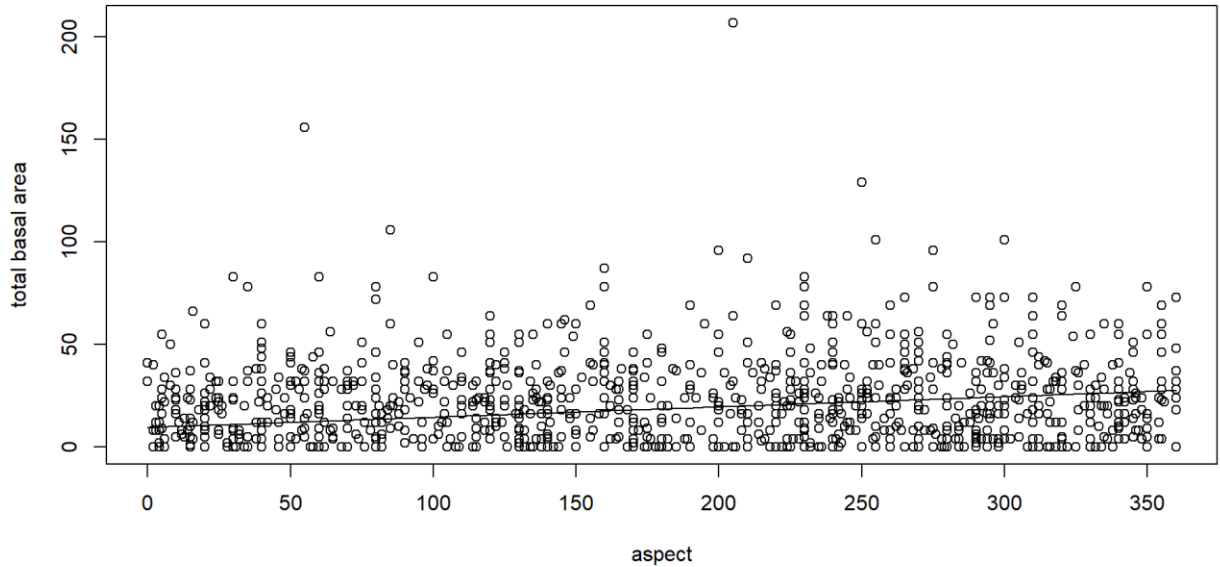


- 1.
2. The histogram is slightly skewed to the left towards the lower elevation range of sampling sites. The highest frequency sites are in the +- 400 meter range. The frequency of sites increases dramatically with elevation from 200-400 meters. Between 4 and 500 meters, frequency of sites decreases sharply, but then levels out with only slight decreases in frequency until around 650 or 700 meters. The majority of the sites fall in the 200-650 meter range.
3. Slope is in units of percent slope (%).
4. There is a fairly even distribution of slope among sampling sites, with more sites that are very flat versus very steep. The highest frequency of slope among the sites are in the 40-70% range. The frequency of sites increases quite evenly with slope from 0-40, then remains high between 40 and 70%, and subsequently decreases sharply after 70%. Most of the sites are in the 10-80% slope range.
5. Aspect is the slope's orientation in degrees, measured counterclockwise from 0-360. 0 degree slopes are north facing, 90 degree slopes are east facing, 180 south facing, 270 west facing.

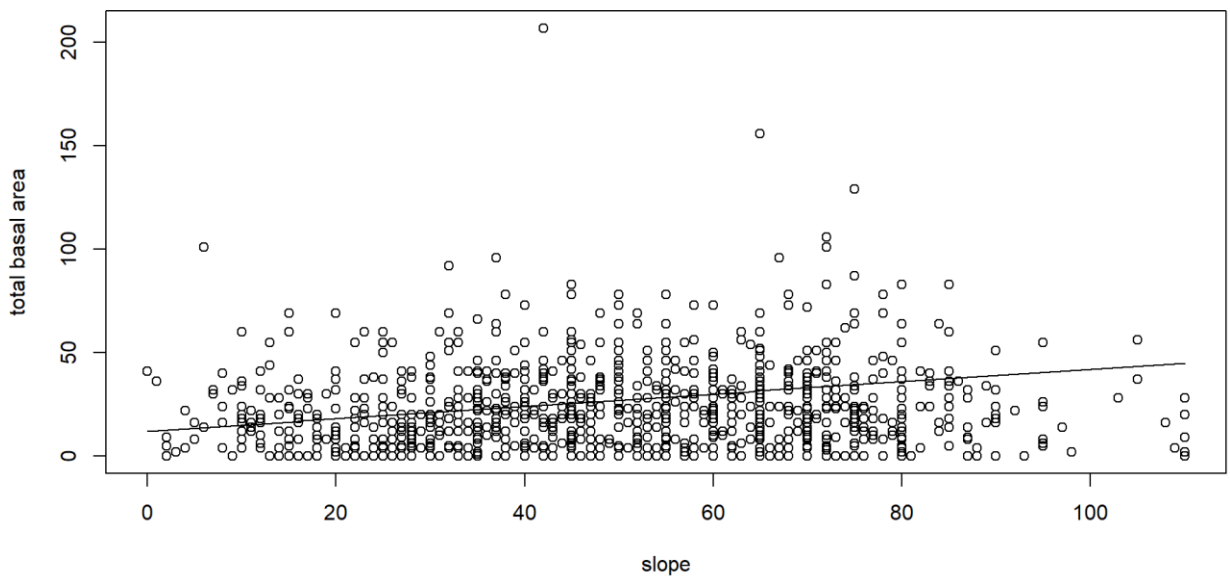
6. The aspect of sampling sites is not evenly distributed, as there is relatively high frequency of sampling sites across all slope orientations. The highest frequency of sampling sites is in the 300 degree range which is west/northwest facing slopes. The lowest frequency of sampling sites were in the 180-200 degree range which is south facing slopes.

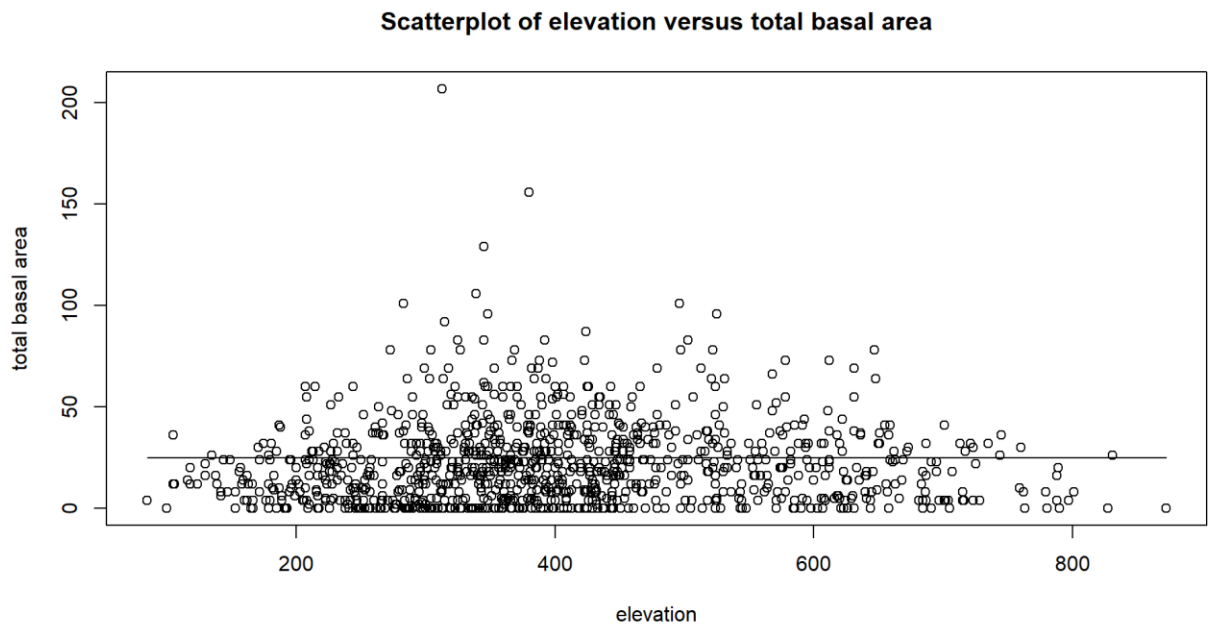
7.

Scatterplot of aspect versus total basal area



Scatterplot of slope versus total basal area





8. For the aspect vs total basal area scatterplot, there is not a very strong association present. There might be a very slight positive relationship between aspect and total basal area, which is described by the line that I fit with a very shallow positive slope. Overall, the linear function does not do a good job of describing the data because there is so much variance that isn't described by a line.

For the slope vs total basal area scatterplot there is very weak association and once again it isn't well described by a linear function. It seems like the data may follow a very slightly parabolic curve, where both very low and very high slopes have low total basal area with some peaks throughout the range of slope values in between.

For the elevation vs total basal area scatter plot, the association seems to be a little more parabolic, where the relationship between elevation and total basal area is positive at first until around 4 or 500 meters and then turns negative between 500 and 800 meters.