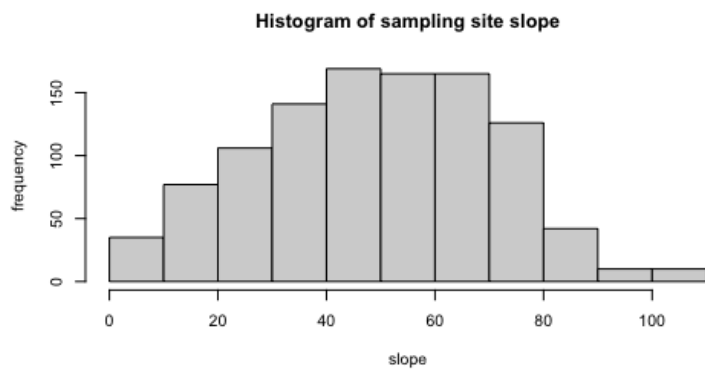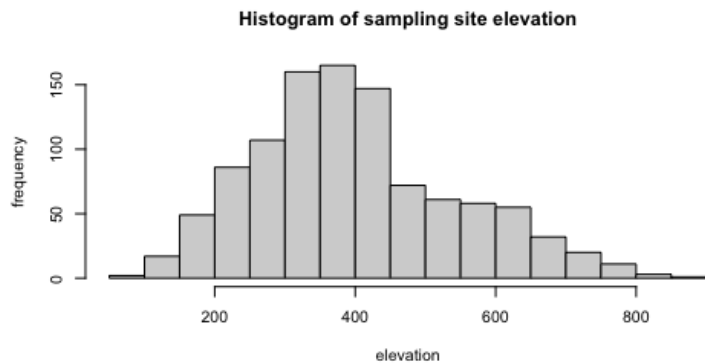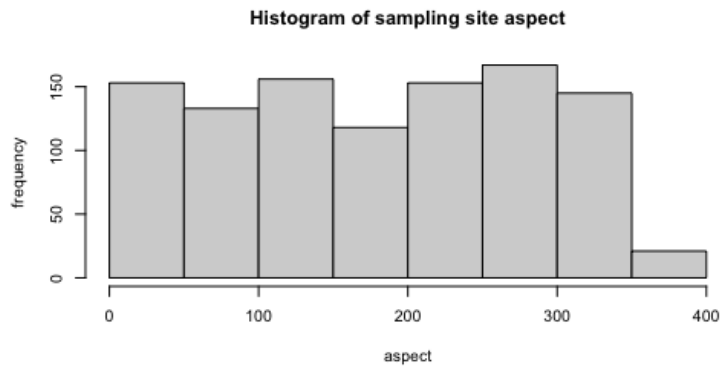Kaitlyn Valentinetti

October 9, 2022

ECo 602: Michael France Nelson

Data Exploration and Deterministic Functions

1.

Histogram of sampling site aspect

Histogram of sampling site elevation

Histogram of sampling site slope

2. The elevation frequency histogram shows that the most sampling sites have an elevation between 200-400 meters. The frequency (number of sampling sites) steadily increases as elevation increases until the most frequent elevation is reached between 350-400m. It's also observable that there is a greater number of low-elevation sampling sites than high-elevation. After the peak at 350-400m, the frequency of elevation steadily declines after

400m. This description is displayed on the graph in the shape of a hill with its peak between 200-400m.

3. The units of slope in this dataset are percent (%).

4. The slope frequency histogram is shaped similarly to the elevation slope histogram because it also has a hill-shaped appearance. In this case, the greatest number of sampling sites have a slope of 40-50%. Similar to the elevation histogram, the slope histogram frequency (number of sampling sites) steadily increases as slope increases until it peaks at 40-50%. Unlike the elevation histogram, the slope histogram does not steadily decrease after the peak. From a slope of 50-70%, the frequency remains constant, and then begins to decrease again. The graph tells us that most of the sampling sites are not flat and have a significantly steep slope.

5. Aspect is the compass orientation of slope that a terrain surface faces in degree units from 0 to 360.

6. The aspect frequency histogram is relatively flat in shape throughout the diagram. This tells us that the frequencies (number of sampling sites) are about the same for all aspect measurements. The lowest frequency by far is at 360 degrees because we must consider that 0 degrees and 360 degrees are the same direction (north). Otherwise, the sampling sites are evenly distributed in all compass directions with the highest frequency between 250-300 degrees (southwest-northwest).

7.

Scatterplot of aspect data

Scatterplot of elevation data

Scatterplot of slope data
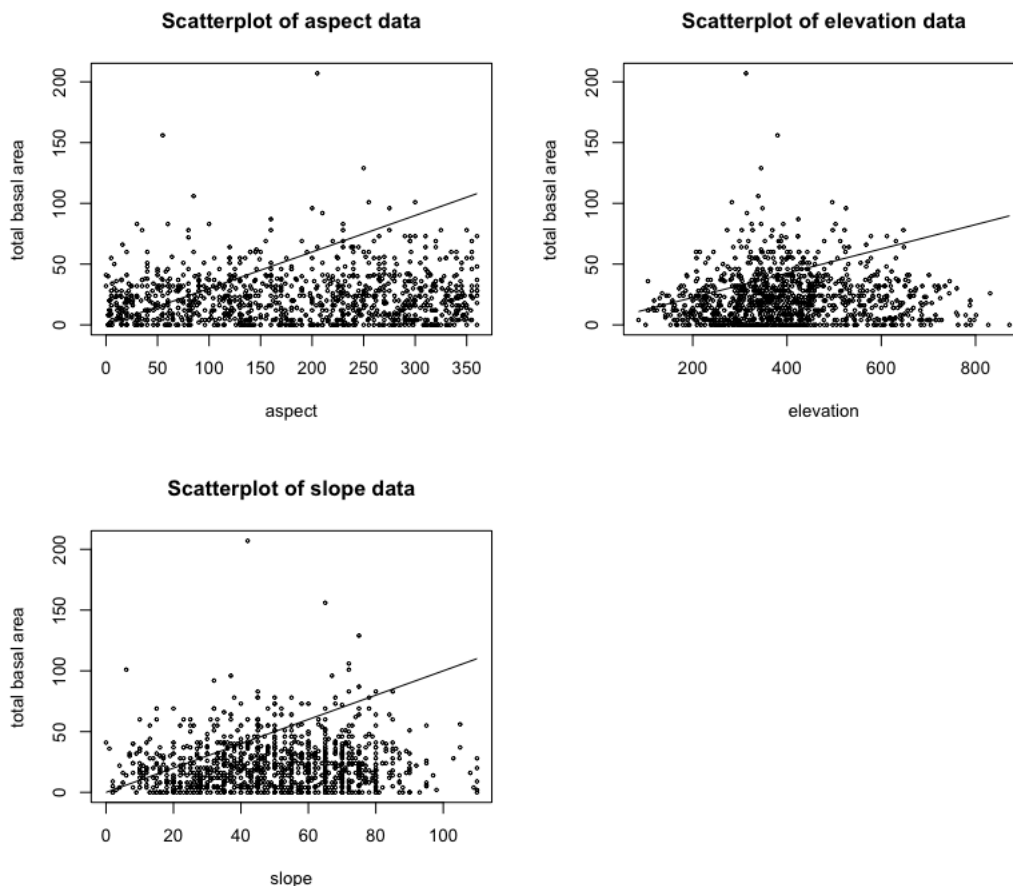
8. In neither elevation, slope, nor aspect is there a clear association between the terrain variable and total basal area. In the aspect scatterplot, most of the data points fall below a basal area of 75m2 per hectare with a few outliers above and are evenly distributed along an aspect gradient. Thus, my linear model is not a great fit for the data because there is no obvious increasing trend in basal area and aspect as the linear model suggests. The elevation scatterplot has a slight positive association between the two variables that is most obvious from 200-400m in elevation. Otherwise, a linear model fit is a bit of a stretch and doesn't fit the entirety of the data. The slope scatterplot arguably has the greatest association between the basal area and terrain variables. The model shows a slightly positive association between basal area and slope, but it is not strong. For this data, I believe a linear model is a relatively good fit for the data because the majority of the points follow along the line.