# MainProjDBH

*PetraGuy*

*11 January 2018*

To run the code the data "table_DBH_live_counts_71-03.csv" is required in the data directory and working directory should be Code.

There are some handy funtions to examine the data: PlotSite(Site) and PlotPlot(Site,Plot) will give bar charts from year 2 aggregated dbh measurements. DBH(Site,Plot) will give the dbh counts for a plot and In(Site,Plot) will give the index for a plot.

## Introduction

The purpose of this exploration is to examine the use of various metrics calculated from the DBH class interval and frequency distribution measurements to generate a metric which correlates with time since disturbance. The successional effect therefore expressed via the metric should correlate with species richness.

... add some stuff in here about stad age, dbh frequency distributions, what they look like etc....

All the dbh data is imported from DBH_live_counts. This is filtered to just look at year 2. Since there are multiple species, each plot has multiple lines per DBH class per plot. These were therefore aggregated, so that this analysis is ignoring species.

The dbh classes in the raw data are recorded in 5cm intervals, class 1 is 5-10cm, class 2 is 5-10cm.

(I haven't found much literature on using DBH in this way. Buchholz, 1978, DBH distribution analysis of different species to infer succession, Lorimar 1983 looked at finding even aged stands, but on looking at dbh frequency distribution says "Weibull shape parameter reflects the degree of skewness, but is not highly sensitive to horizontal displacement of mode from median", he uses a symmetry index which depends on the mode,and therefore only useful for unimodal data)

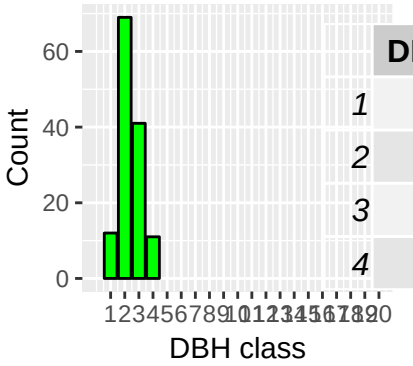
## Running sum of differences

First, the sum of running differences in count between adjacent DBH classes is trialled. Using this calculation it can be seen that a high frequency of small dbh classes would be the first term in the sum. Subsequent subtractions of smaller values would result in a large positive value. A large positive would then represent an early succession with lots of young saplings. The reverse would occur for a mature plot and a large negative number would be found.

```
## Joining, by = c("PLOT", "SITE")
```

Consider site 2 plots 10 and 11

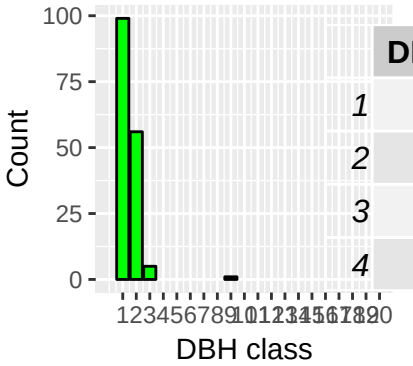
site 2 plot 10




| | DBH_class | PLOT | SITE | Count |
|---|---|---|---|---|
| 1 | 1 | 10 | 2 | 12 |
| 2 | 2 | 10 | 2 | 69 |
| 3 | 3 | 10 | 2 | 41 |
| 4 | 4 | 10 | 2 | 11 |

| | PLOT | SITE | INDEX |
|---|---|---|---|
| 1 | 10 | 2 | −109 |

site 2 plot 11



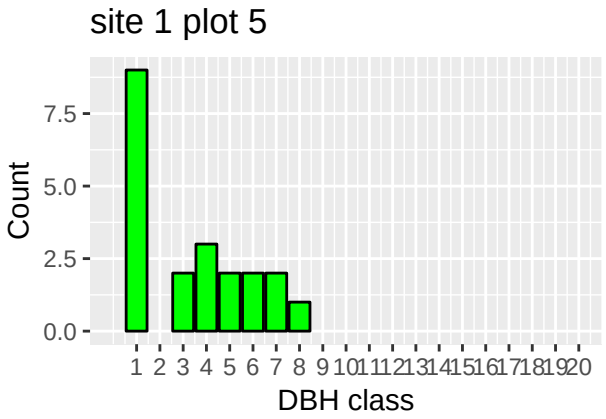

| | DBH_class | PLOT | SITE | Count |
|---|---|---|---|---|
| 1 | 1 | 11 | 2 | 99 |
| 2 | 2 | 11 | 2 | 56 |
| 3 | 3 | 11 | 2 | 5 |
| 4 | 9 | 11 | 2 | 1 |

| | PLOT | SITE | INDEX |
|---|---|---|---|
| 1 | 11 | 2 | 37 |

Plots 10 and 11 have lots of smaller trees, and therefore a similar time since an opening event. But they have indices of - 109 and 37 respectively. Plot 10 site 2 has the most negative index of all the plots, so we would want this number to imply a long time has occurred since an opening event, which it does not.

Looking at the counts, plot 10 has 12 in the first category and 69 in the next, therefore the first subtraction gives a relatively large negative value, the remaining values, all being smaller, result in a large negative index.

A negative can occur in a right skew distribution if the count of trees in the first class is low, for example, Site 1 Plot 5

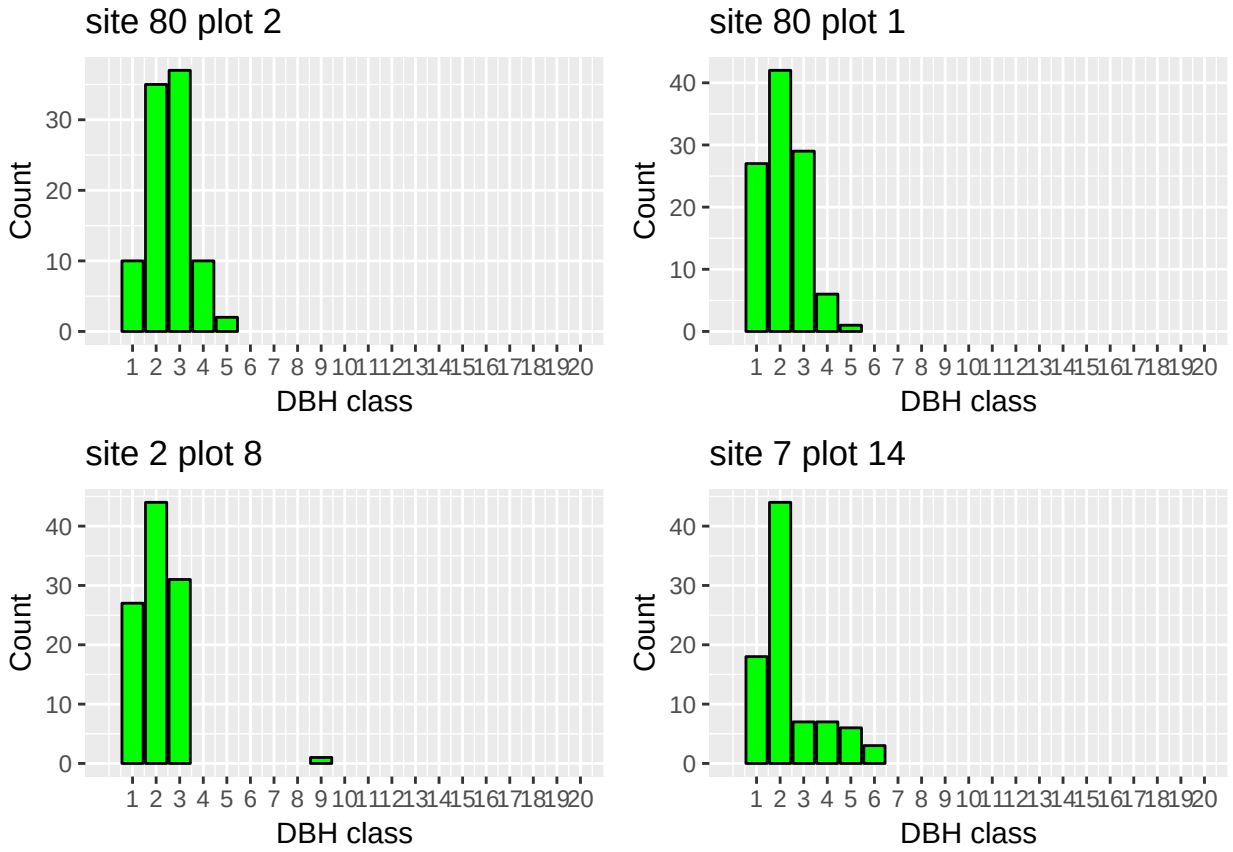



| | DBH_class | PLOT | SITE | Count |
|---|---|---|---|---|
| *1* | 1 | 5 | 1 | 9 |
| *2* | 3 | 5 | 1 | 2 |
| *3* | 4 | 5 | 1 | 3 |
| *4* | 5 | 5 | 1 | 2 |
| *5* | 6 | 5 | 1 | 2 |
| *6* | 7 | 5 | 1 | 2 |
| *7* | 8 | 5 | 1 | 1 |

| | PLOT | SITE | INDEX |
|---|---|---|---|
| *1* | 5 | 1 | −3 |

This shows that the subtraction works if there are successively fewer counts in each class, and if the sum of the counts in the classes above class 1 are less than those in the first class. The most negative indices occur for sites with a large number of trees in dbh class 2, which is greater than, or similar to the number in class 1.

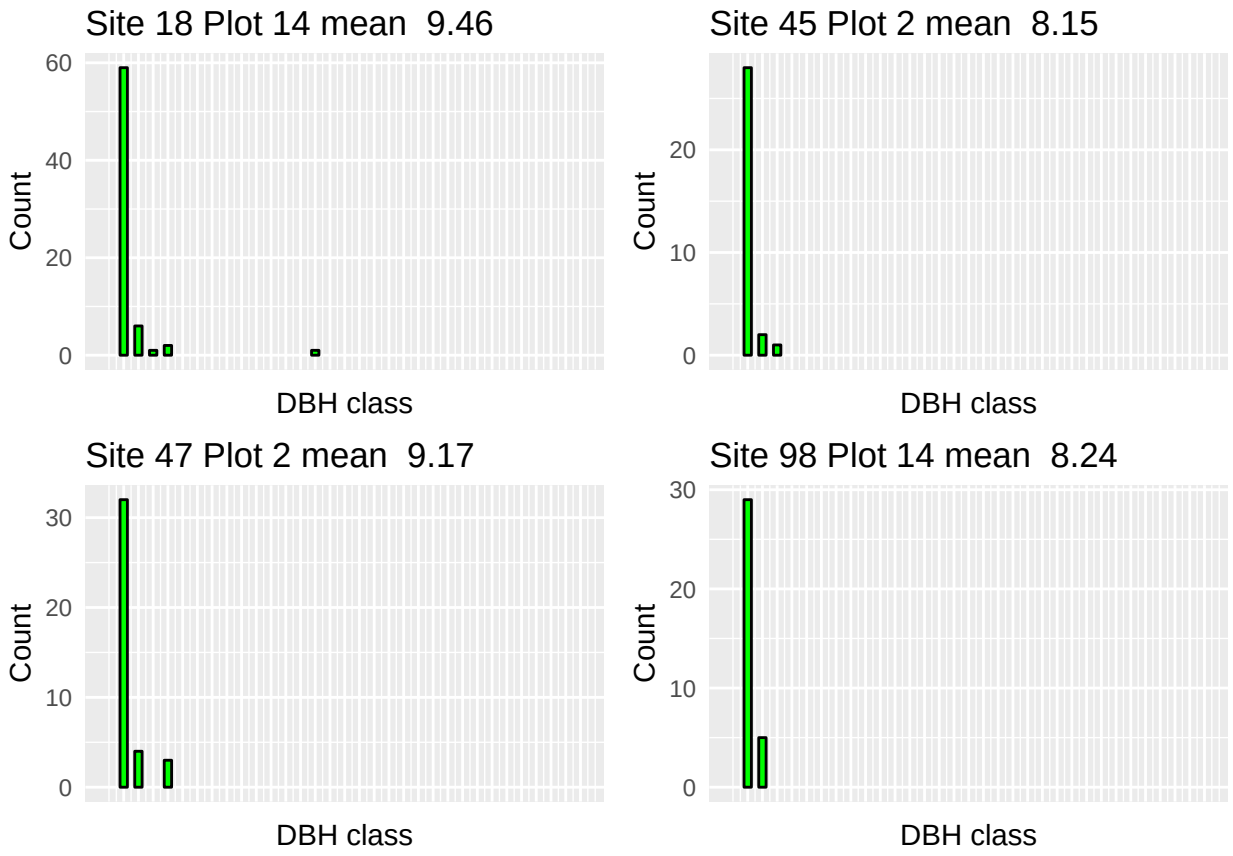
```
##     DBH_class PLOT SITE Count INDEX
## 10          1   10    2    12  -109
## 293         1    2   80    10   -74
## 292         1    1   80    27   -51
## 8           1    8    2    27   -49
## 39          1   14    7    18   -49
## 156         1   12   37     2   -45
```



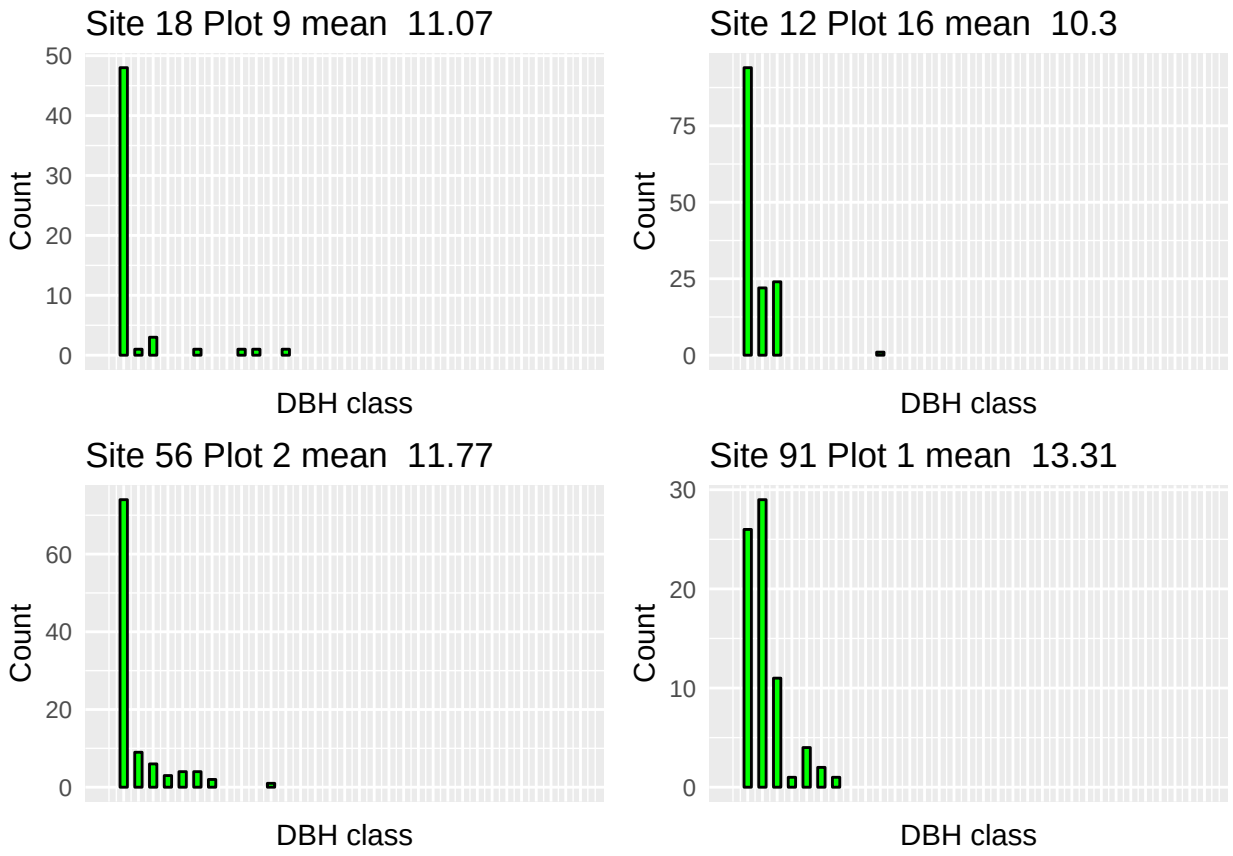

The above graphs show the plots with the four most negative indices. In each case it can be seen that combining the classes would resolve the problem. It could be argued that this is a valid solution because we are looking for a large number of small trees of a similar size, and for that size to drop of rapidly. However, since the histogram should represent the underlying frequency distribution, then these larger groups are loosing information and are not representative of the data. Therefore a solution that uses the small bin widths of the original data is preferable.

## The mean as a metric

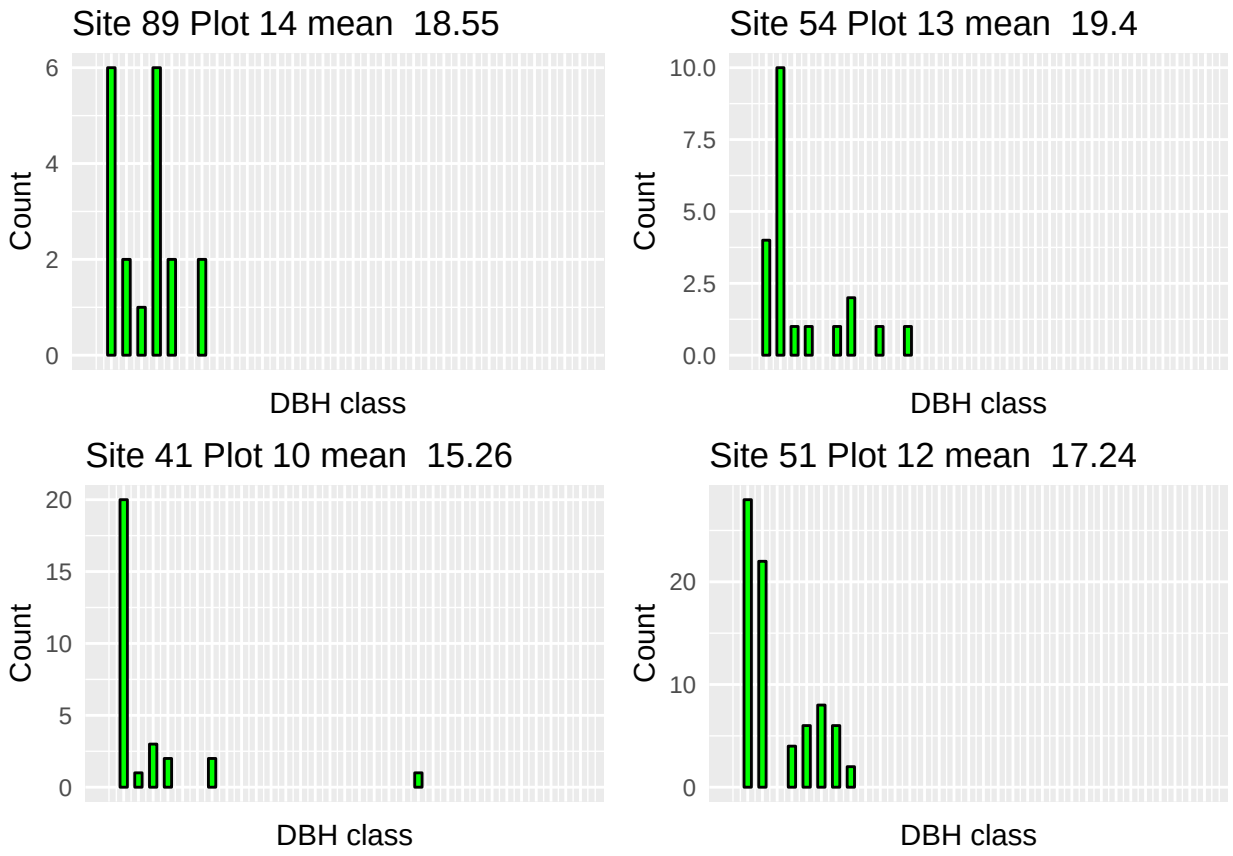
Plots of plots with mean less than 10



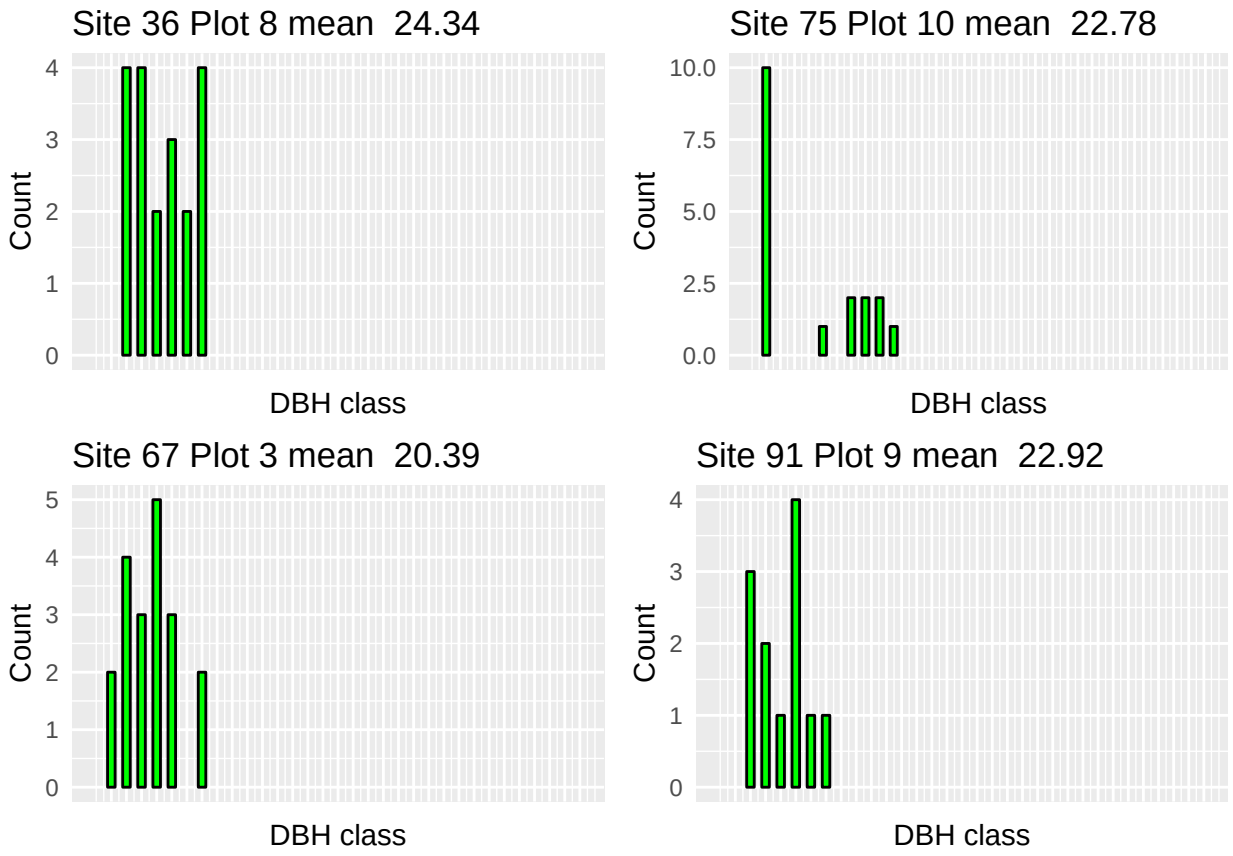

For a mean below 10 the graphs demonstrate a pattern consistent with lots of young saplings and a recent opening event.



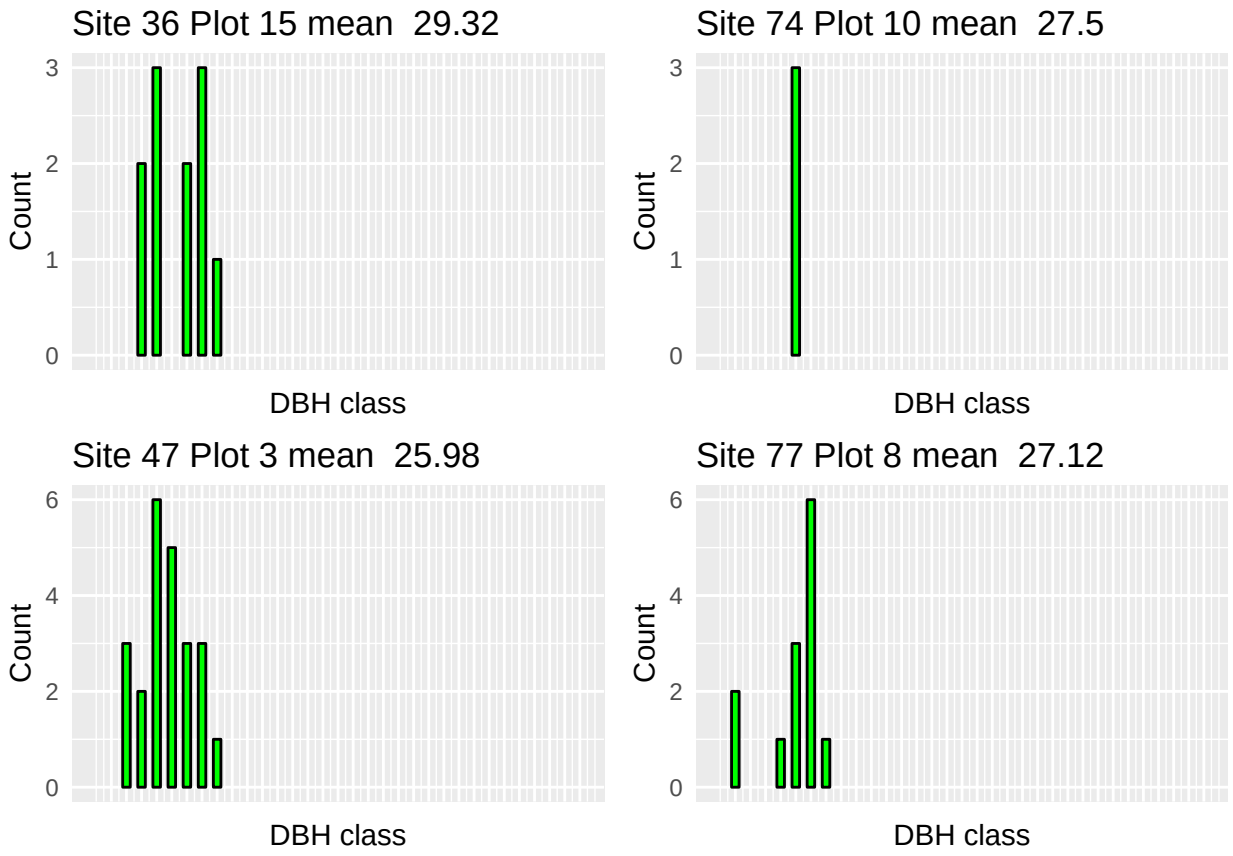

For means between 10 and 15 the frequency of trees in the next DBH classes increases



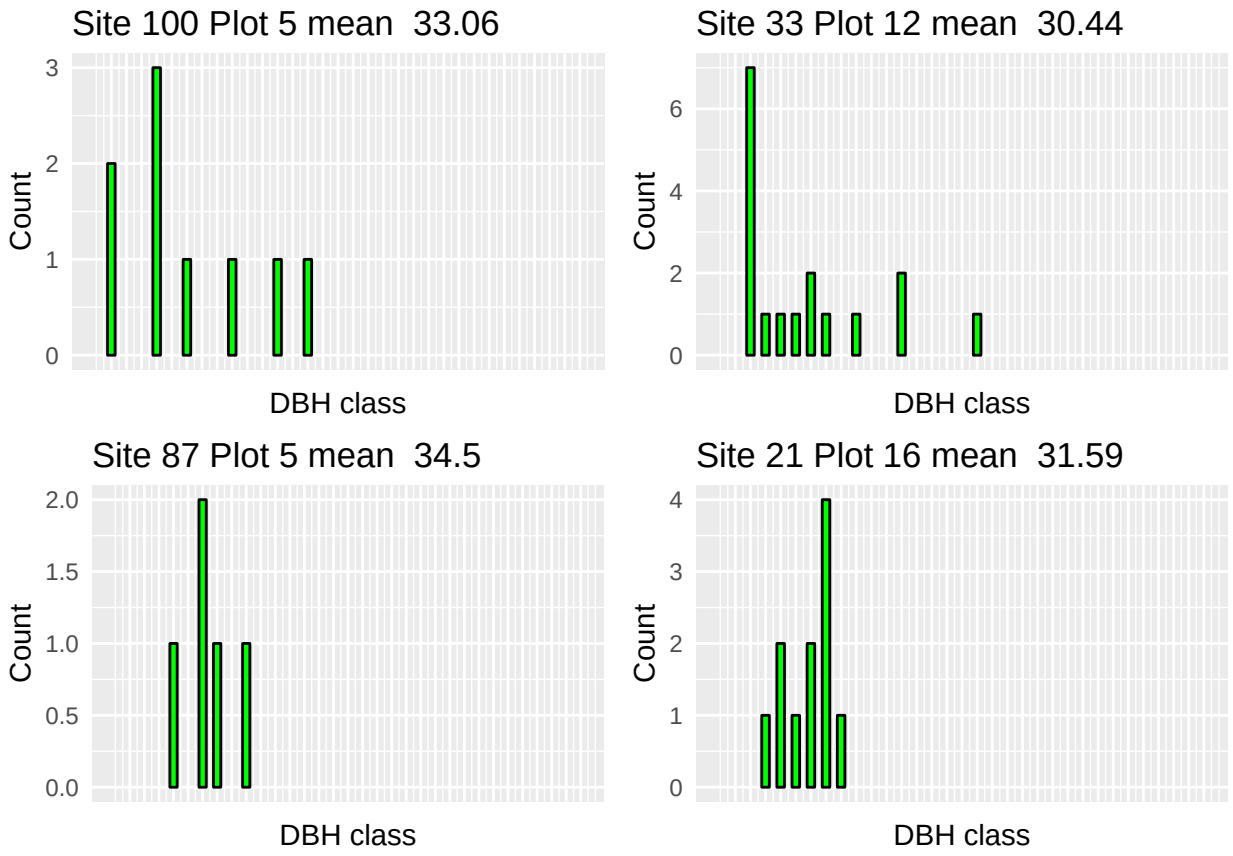

As the mean increases the counts in the larger DBH classes increases, but the distribution still resembles early succession.




Means over 20 look like later successional distributions

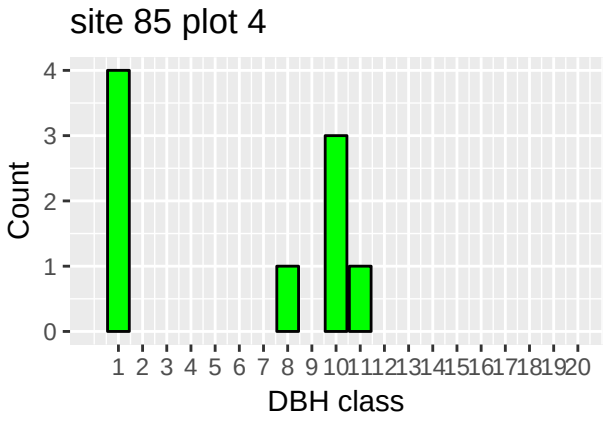



The means over 25 are varied, Site 18 plot 6 loks like early succession, wheres all the others do not.

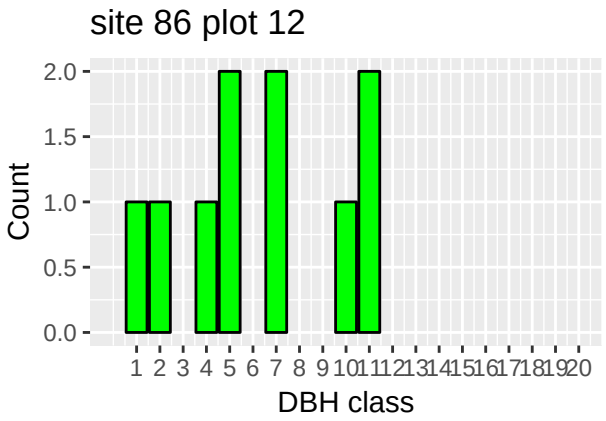



Means over 30 appear to represent later succession more frequently.

With low means, because the mean is in or near the lowest dbh class, the distribution will be right skewed and fit the desired criteria of representing a plot with lots of young trees and few older ones. When we consider larger means there are more distributions that could give that mean, and whether or not they reflect the time since an opening event is less clear. Consider Site 85 plot 4 and Site 86 plot 12.




31.94




34

These have similar means, but would we be able to say they represented plots that had similar time since event. Site 85 plot 4 with its righ skew and some saplings could be in an early successional stage, whereas Site 86 plot 12 looks more mature.

The mean may not be a metric that clearly correlates with time since event.

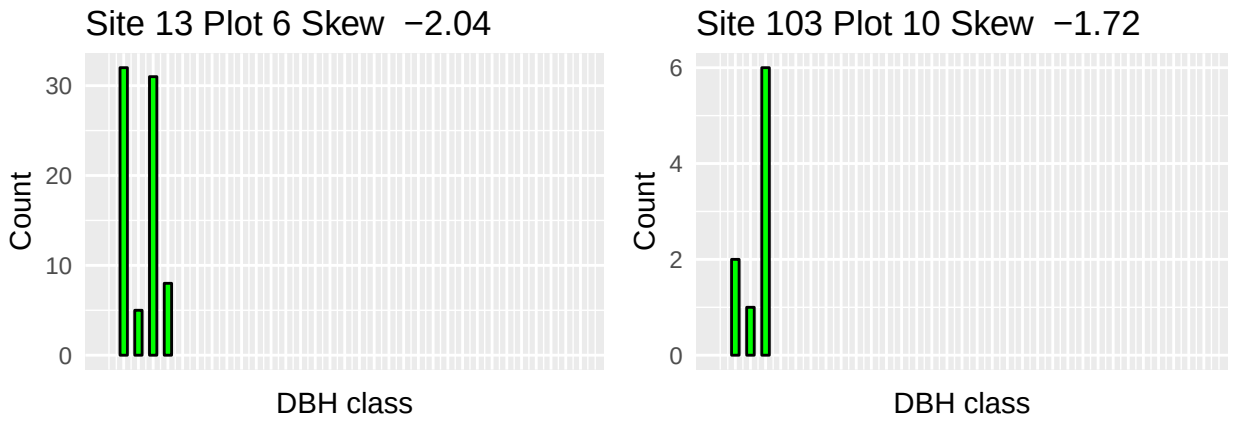
## Skew as a metric.

We could use Pearson's second skewness coefficient:

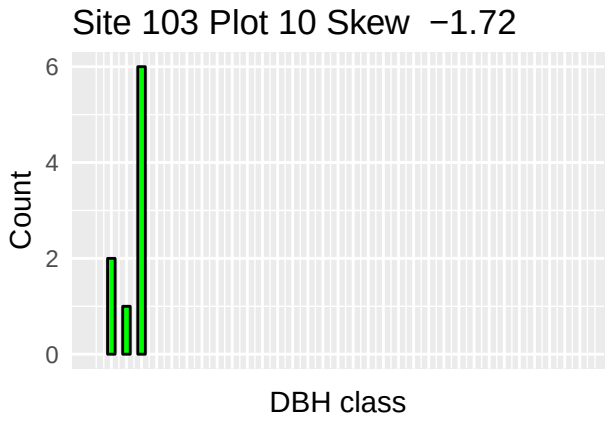
$$Skew = \frac{3(mean - median)}{s}$$

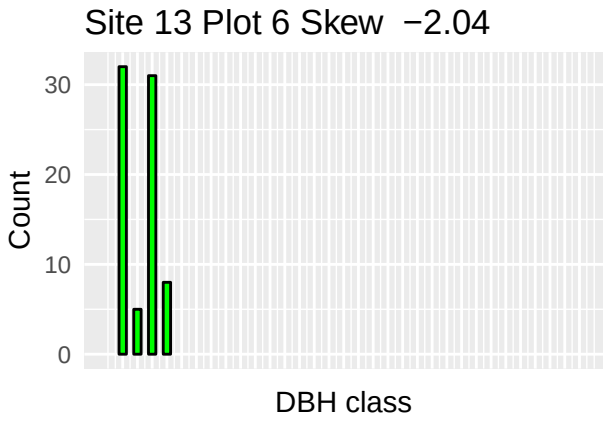
Or Pearson's third moment

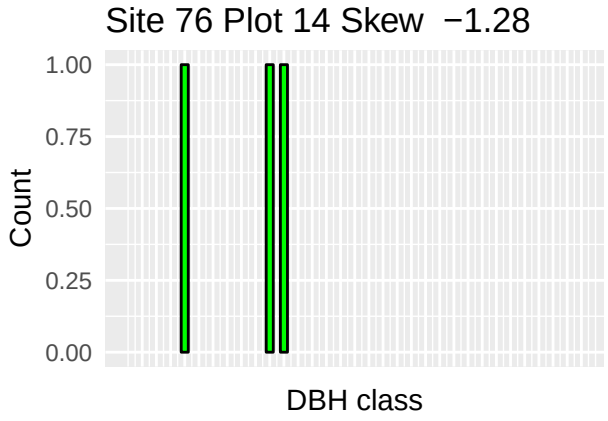
$$Skew_3 = \frac{\frac{\Sigma(x_i - mean)^3}{n}}{S^3}$$

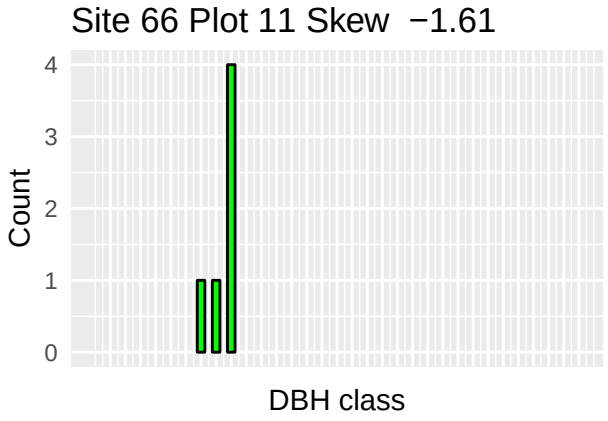

Site 13 Plot 6 Skew −2.04
Count
DBH class

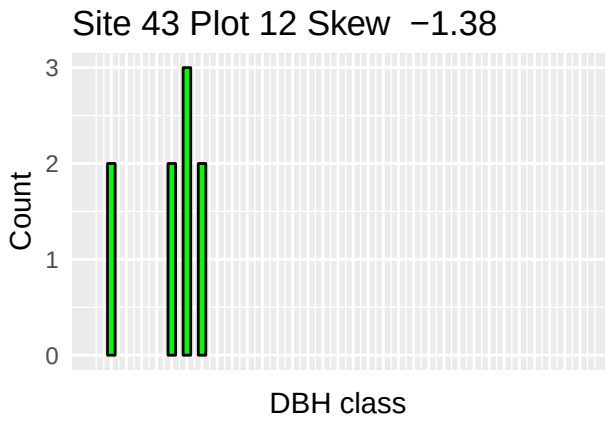

Site 103 Plot 10 Skew −1.72
Count
DBH class

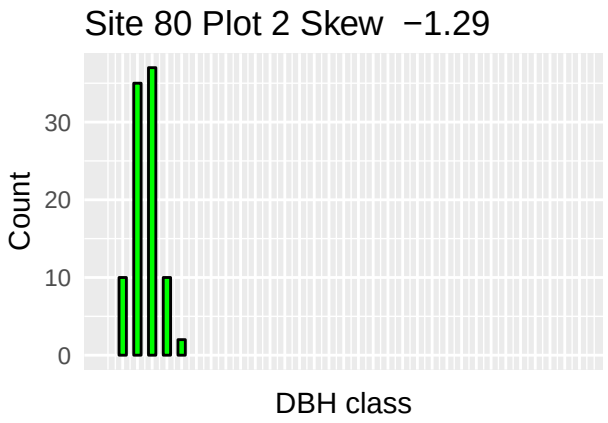

Site 103 Plot 10 Skew −1.72
Count
DBH class

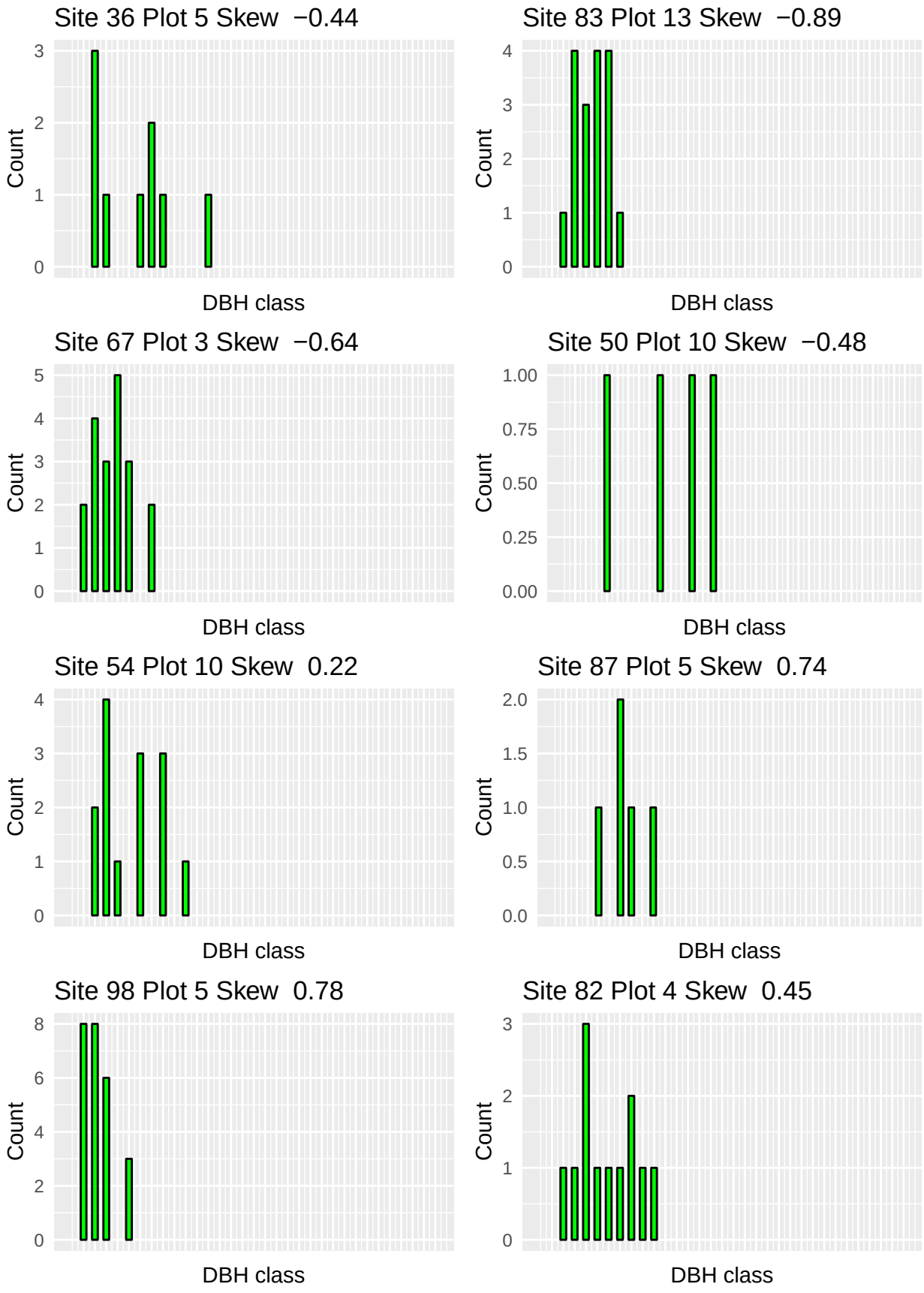

Site 13 Plot 6 Skew −2.04
Count
DBH class

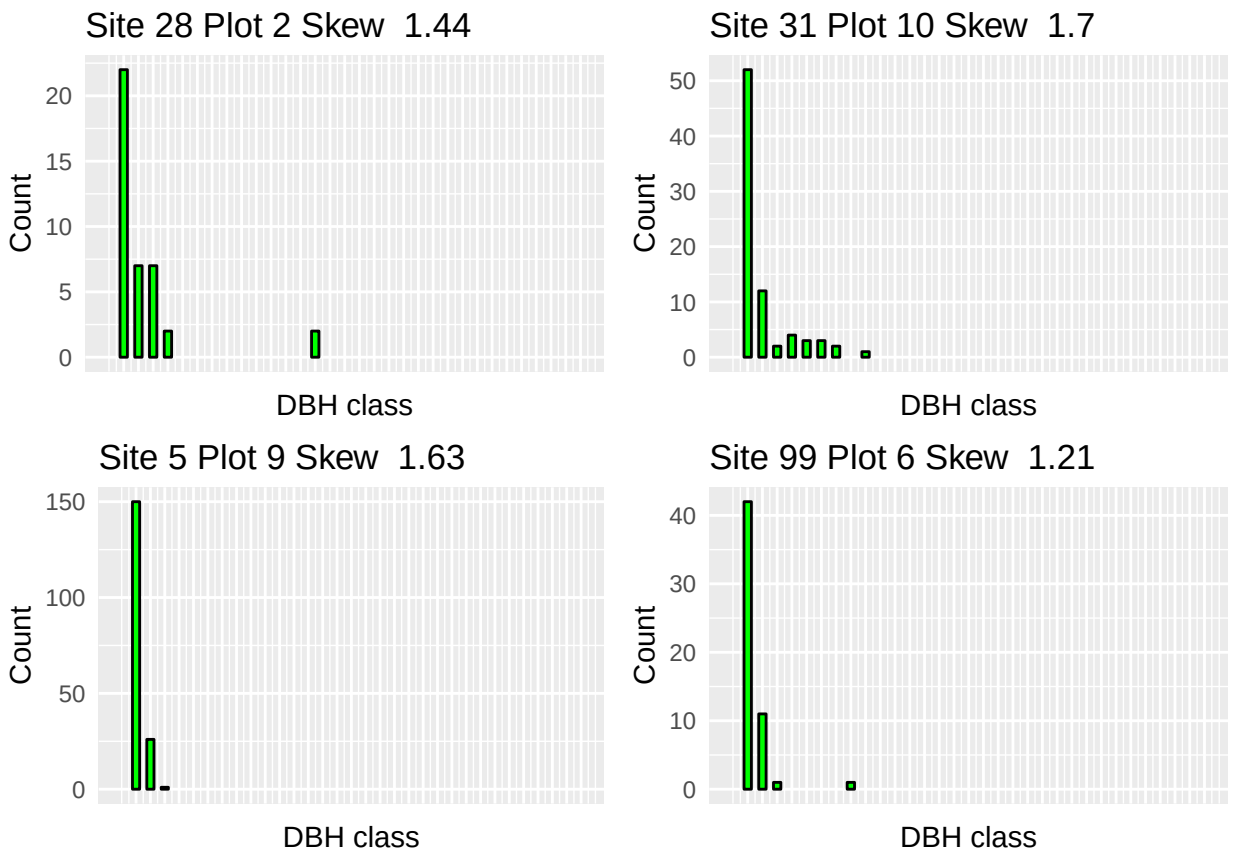

Site 76 Plot 14 Skew −1.28
Count
DBH class


Site 66 Plot 11 Skew −1.61
Count
DBH class

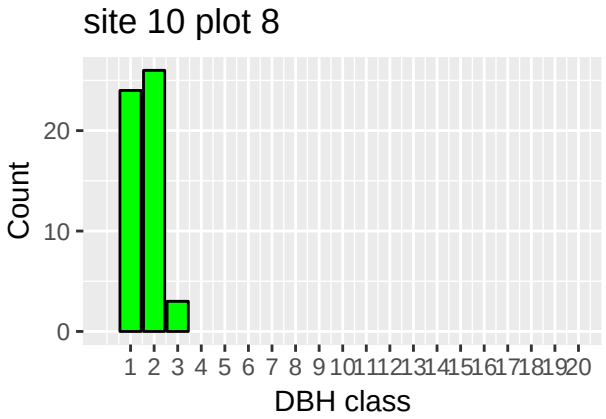

Site 43 Plot 12 Skew −1.38
Count
DBH class


Site 80 Plot 2 Skew −1.29
Count
DBH class

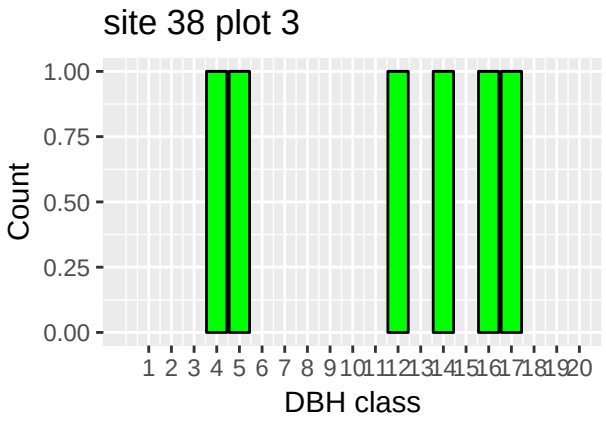

Site 36 Plot 5 Skew −0.44
Site 83 Plot 13 Skew −0.89
Site 67 Plot 3 Skew −0.64
Site 50 Plot 10 Skew −0.48
Site 54 Plot 10 Skew 0.22
Site 87 Plot 5 Skew 0.74
Site 98 Plot 5 Skew 0.78
Site 82 Plot 4 Skew 0.45
Count
DBH class




The most positively skewed data accurately represents the sites with the most counts in the lowet dbh class, however, as the skew becomes more negative a pattern of progression to more negatively skewed distributions and hence more later successional plots is not clear. For example, compare Site 10 Plot 8 and Site 38 plot 3



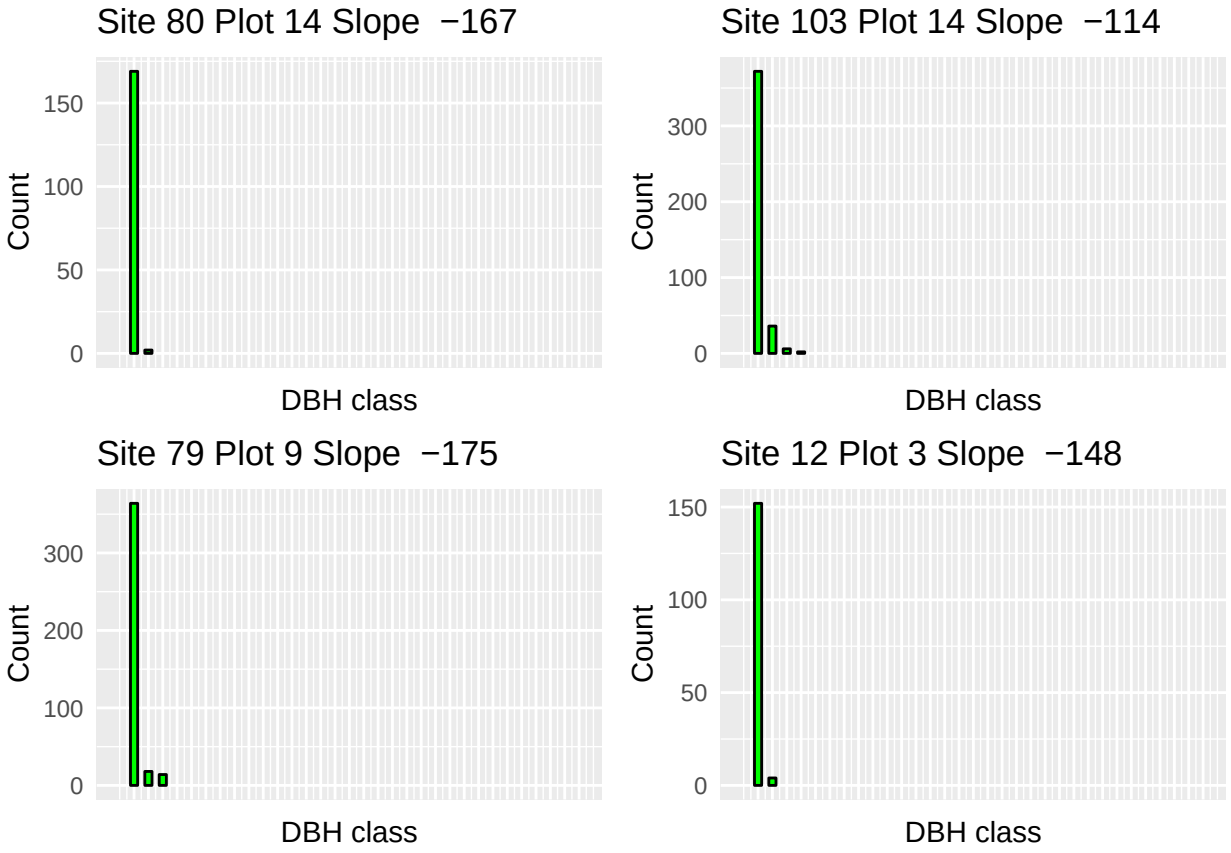

| | mean1 | med1 | skew1 |
|---|---|---|---|
| *1* | 10.52 | 12.25 | −1.73 |

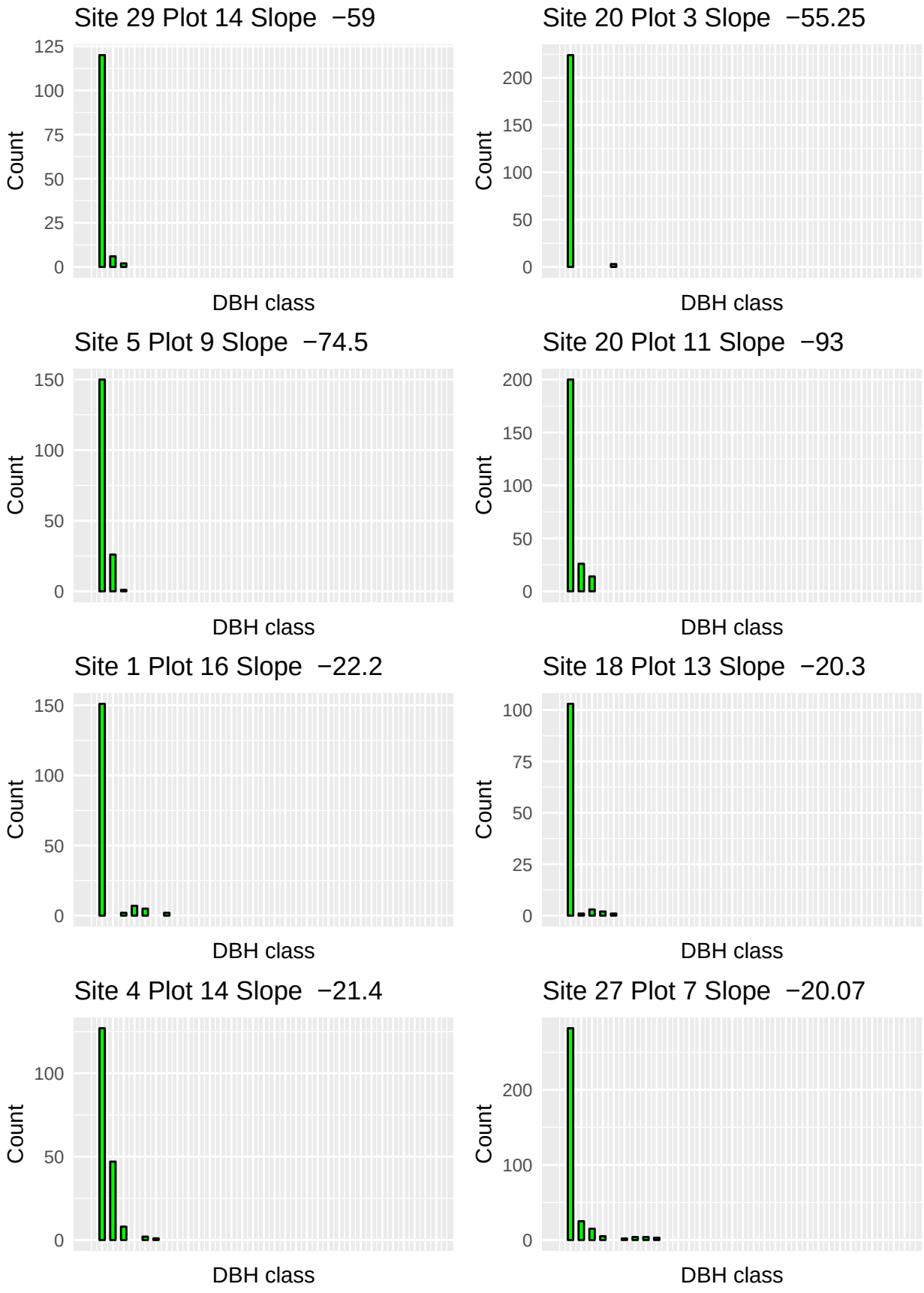


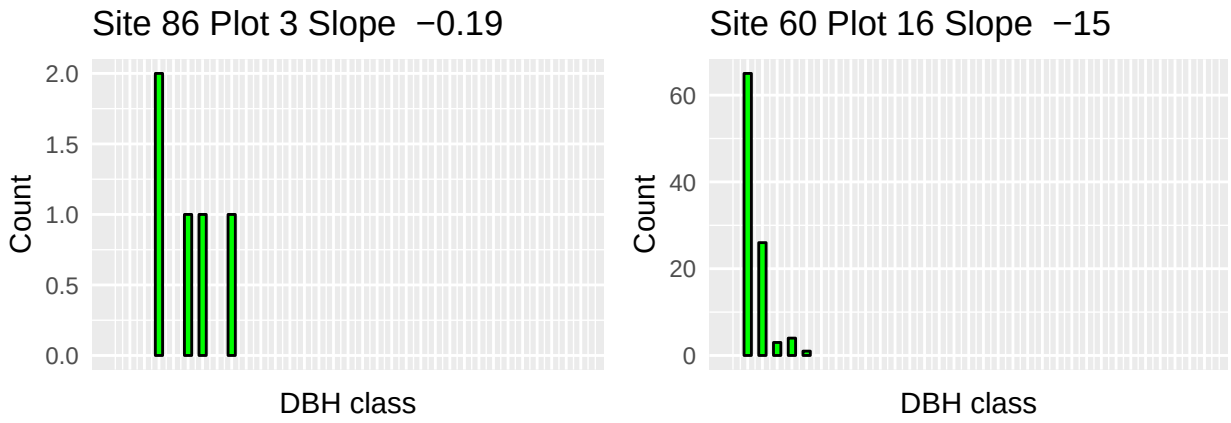

| | mean2 | med2 | skew2 |
|---|---|---|---|
| *1* | 59.17 | 67.25 | −0.87 |

The distributions are very different representing very different plots. The first plot has a lot of saplings and we would want this to be showing a positive skew. The skew calculation give a negative value because the large number of small trees give a small mean, but the slightly larger count in dbh class 2 results in a median greater than the mean. For Site 38 plot 3 we would like to seeing a skew of zero representing a few large trees and no saplings.

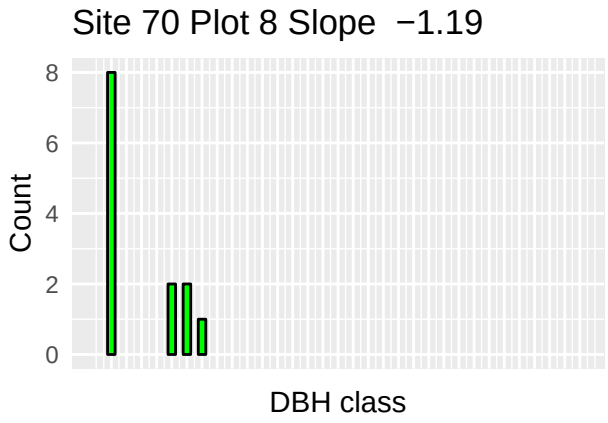
**Regression fit.**

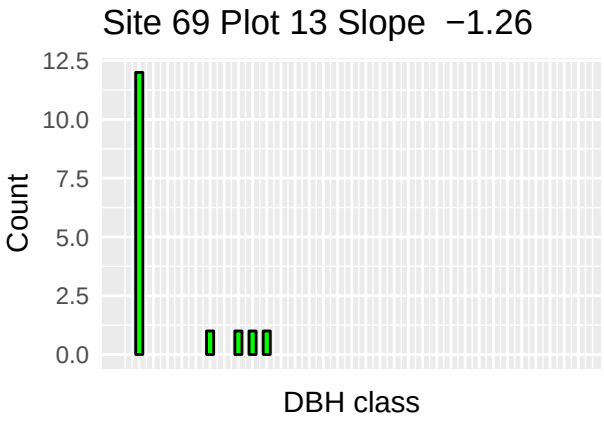


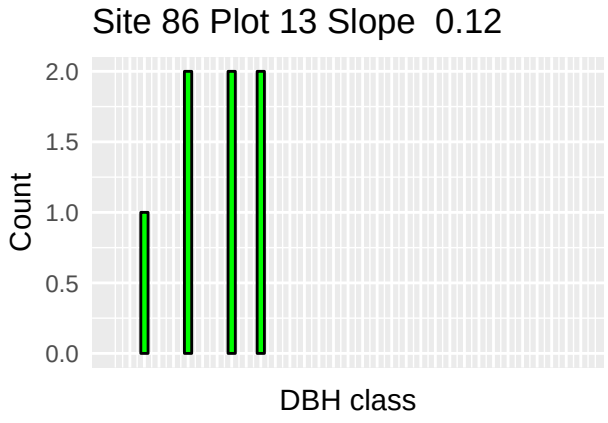

Site 29 Plot 14 Slope −59
Site 20 Plot 3 Slope −55.25
Site 5 Plot 9 Slope −74.5
Site 20 Plot 11 Slope −93
Site 1 Plot 16 Slope −22.2
Site 18 Plot 13 Slope −20.3
Site 4 Plot 14 Slope −21.4
Site 27 Plot 7 Slope −20.07
Count
DBH class

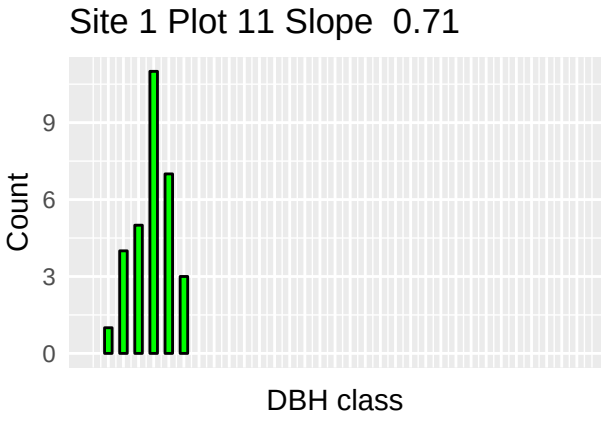

Site 86 Plot 3 Slope −0.19
Count
DBH class

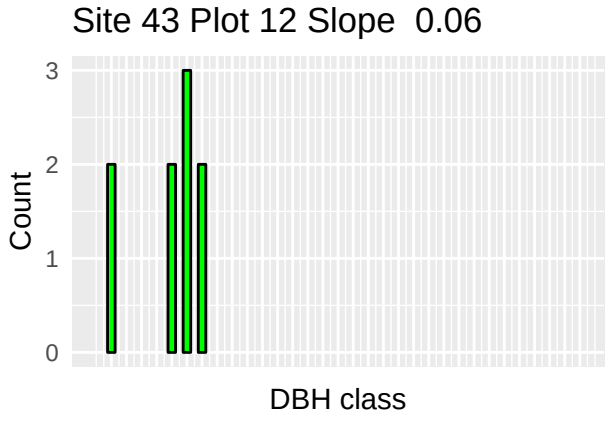

Site 60 Plot 16 Slope −15
Count
DBH class

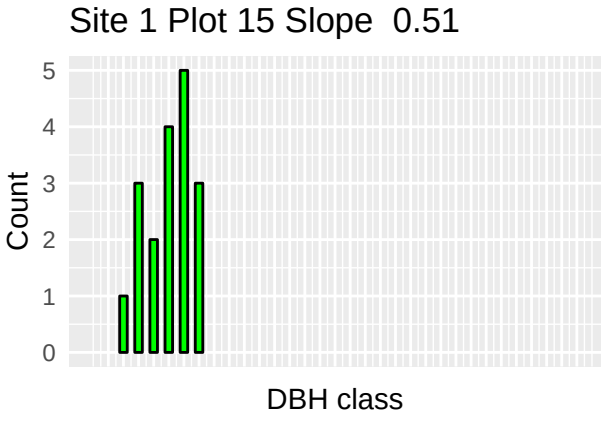

Site 70 Plot 8 Slope −1.19
Count
DBH class

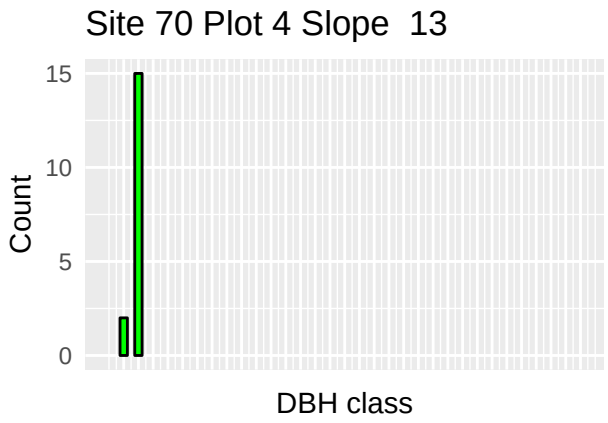

Site 69 Plot 13 Slope −1.26
Count
DBH class

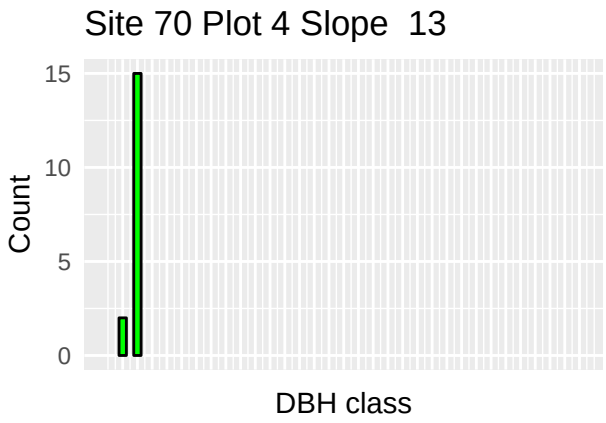

Site 86 Plot 13 Slope 0.12
Count
DBH class

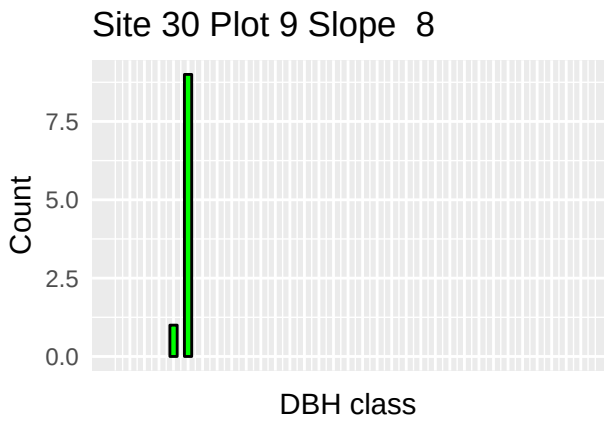

Site 1 Plot 11 Slope 0.71
Count
DBH class

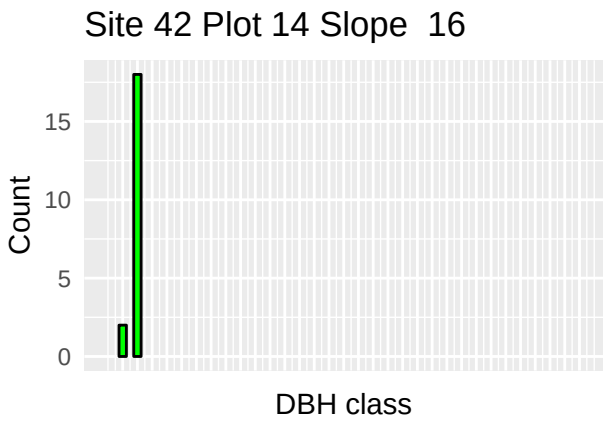

Site 43 Plot 12 Slope 0.06
Count
DBH class


Site 1 Plot 15 Slope 0.51
Count
DBH class


Site 70 Plot 4 Slope 13
Count
15
10
5
0
DBH class


Site 70 Plot 4 Slope 13
Count
15
10
5
0
DBH class


Site 30 Plot 9 Slope 8
Count
7.5
5.0
2.5
0.0
DBH class


Site 42 Plot 14 Slope 16
Count
15
10
5
0
DBH class