

# MainProjSiteDescriptorAnalysis

*PetraGuy*

*8 January 2018*

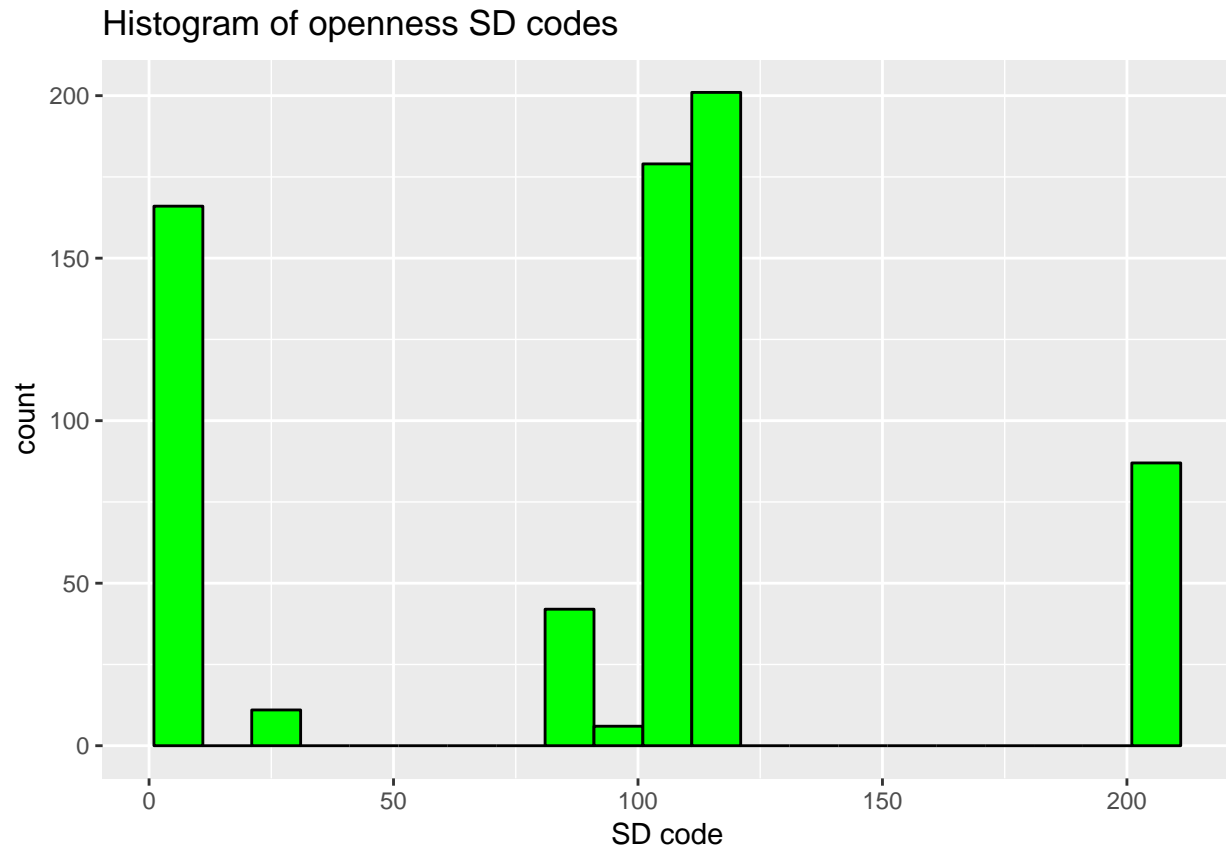
This analysis is to examine site descriptors, to get a feel for which ones are used and how they might be used as part of a “heterogeneity” index or whether they need to be part of any other covariates important for the analysis of species richness and species area relationships.

Codes have been selected that relate to open areas, large ponds, rivers, moorland, rough pasture etc. The ponds and rivers were selected, not because they represent a different habitat, but because they are likely to imply open ground around them.

```
OpenCodesSites = c(8,9,10,24,86,87,90,91,92,93,105,106,107,108,109,110,111,112,113,114,115,116,117,208,118,119,120,121,122,123,124,125,126,127,128,129,130,131,132,133,134,135,136,137,138,139,140,141,142,143,144,145,146,147,148,149,150,151,152,153,154,155,156,157,158,159,160,161,162,163,164,165,166,167,168,169,170,171,172,173,174,175,176,177,178,179,180,181,182,183,184,185,186,187,188,189,190,191,192,193,194,195,196,197,198,199,200,201,202,203,204,205,206,207,209,210,211,212,213,214,215,216,217,218,219,220,221,222,223,224,225,226,227,228,229,230,231,232,233,234,235,236,237,238,239,240,241,242,243,244,245,246,247,248,249,250,251,252,253,254,255,256,257,258,259,260,261,262,263,264,265,266,267,268,269,270,271,272,273,274,275,276,277,278,279,280,281,282,283,284,285,286,287,288,289,290,291,292,293,294,295,296,297,298,299,300,301,302,303,304,305,306,307,308,309,310,311,312,313,314,315,316,317,318,319,320,321,322,323,324,325,326,327,328,329,330,331,332,333,334,335,336,337,338,339,340,341,342,343,344,345,346,347,348,349,350,351,352,353,354,355,356,357,358,359,360,361,362,363,364,365,366,367,368,369,370,371,372,373,374,375,376,377,378,379,380,381,382,383,384,385,386,387,388,389,390,391,392,393,394,395,396,397,398,399,400,401,402,403,404,405,406,407,408,409,410,411,412,413,414,415,416,417,418,419,420,421,422,423,424,425,426,427,428,429,430,431,432,433,434,435,436,437,438,439,440,441,442,443,444,445,446,447,448,449,450,451,452,453,454,455,456,457,458,459,460,461,462,463,464,465,466,467,468,469,470,471,472,473,474,475,476,477,478,479,480,481,482,483,484,485,486,487,488,489,490,491,492,493,494,495,496,497,498,499,500,501,502,503,504,505,506,507,508,509,510,511,512,513,514,515,516,517,518,519,520,521,522,523,524,525,526,527,528,529,530,531,532,533,534,535,536,537,538,539,540,541,542,543,544,545,546,547,548,549,550,551,552,553,554,555,556,557,558,559,560,561,562,563,564,565,566,567,568,569,570,571,572,573,574,575,576,577,578,579,580,581,582,583,584,585,586,587,588,589,590,591,592,593,594,595,596,597,598,599,600,601,602,603,604,605,606,607,608,609,610,611,612,613,614,615,616,617,618,619,620,621,622,623,624,625,626,627,628,629,630,631,632,633,634,635,636,637,638,639,640,641,642,643,644,645,646,647,648,649,650,651,652,653,654,655,656,657,658,659,660,661,662,663,664,665,666,667,668,669,670,671,672,673,674,675,676,677,678,679,680,681,682,683,684,685,686,687,688,689,690,691,692,693,694,695,696,697,698,699,700,701,702,703,704,705,706,707,708,709,710,711,712,713,714,715,716,717,718,719,720,721,722,723,724,725,726,727,728,729,730,731,732,733,734,735,736,737,738,739,740,741,742,743,744,745,746,747,748,749,750,751,752,753,754,755,756,757,758,759,760,761,762,763,764,765,766,767,768,769,770,771,772,773,774,775,776,777,778,779,780,781,782,783,784,785,786,787,788,789,790,791,792,793,794,795,796,797,798,799,800,801,802,803,804,805,806,807,808,809,810,811,812,813,814,815,816,817,818,819,820,821,822,823,824,825,826,827,828,829,830,831,832,833,834,835,836,837,838,839,840,841,842,843,844,845,846,847,848,849,850,851,852,853,854,855,856,857,858,859,860,861,862,863,864,865,866,867,868,869,870,871,872,873,874,875,876,877,878,879,880,881,882,883,884,885,886,887,888,889,890,891,892,893,894,895,896,897,898,899,900,901,902,903,904,905,906,907,908,909,910,911,912,913,914,915,916,917,918,919,920,921,922,923,924,925,926,927,928,929,930,931,932,933,934,935,936,937,938,939,940,941,942,943,944,945,946,947,948,949,950,951,952,953,954,955,956,957,958,959,960,961,962,963,964,965,966,967,968,969,970,971,972,973,974,975,976,977,978,979,980,981,982,983,984,985,986,987,988,989,990,991,992,993,994,995,996,997,998,999,1000)
```

```
## # A tibble: 696 x 2
##   SITE SD_code
##   <int> <int>
## 1     3    105
## 2     3    107
## 3     3    108
## 4     3    114
## 5     3    116
## 6     1     8
## 7     1     9
## 8     1    114
## 9     1    115
## 10    1    117
## # ... with 686 more rows
```

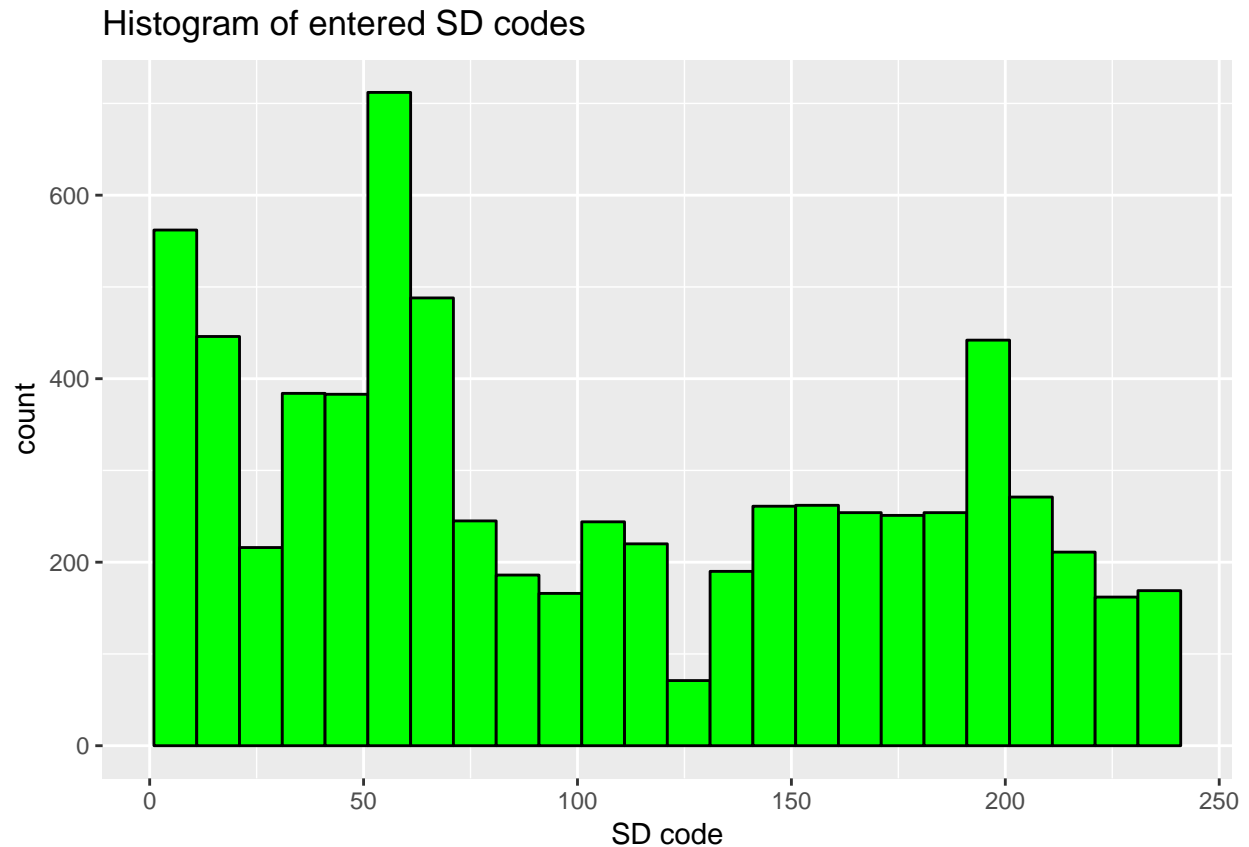
```
ggplot(OpenSites, aes(x = SD_code)) +
  geom_histogram(breaks = seq(1,220, by = 10), col = "black", fill = "green")+
  labs(title = "Histogram of openness SD codes", x = "SD code")
```



The histogram shows that there are many records that could be considered as contributing toward open areas in woodlands, the 0-10 codes are for coppicing, 100 -120 codes are glades, rides, fields, paths, marshes and ditches. Those above 200 are for rough grassland, rivers, lakes and heaths.

Plot below shows the frequency of all the SD codes for comparison

```
ggplot(SiteDescriptorsYr2, aes(SiteDescriptorsYr2$SD_code)) +
  geom_histogram(breaks = seq(1,250, by = 10), col = "black", fill = "green")+
  labs(title = "Histogram of entered SD codes", x = "SD code")
```



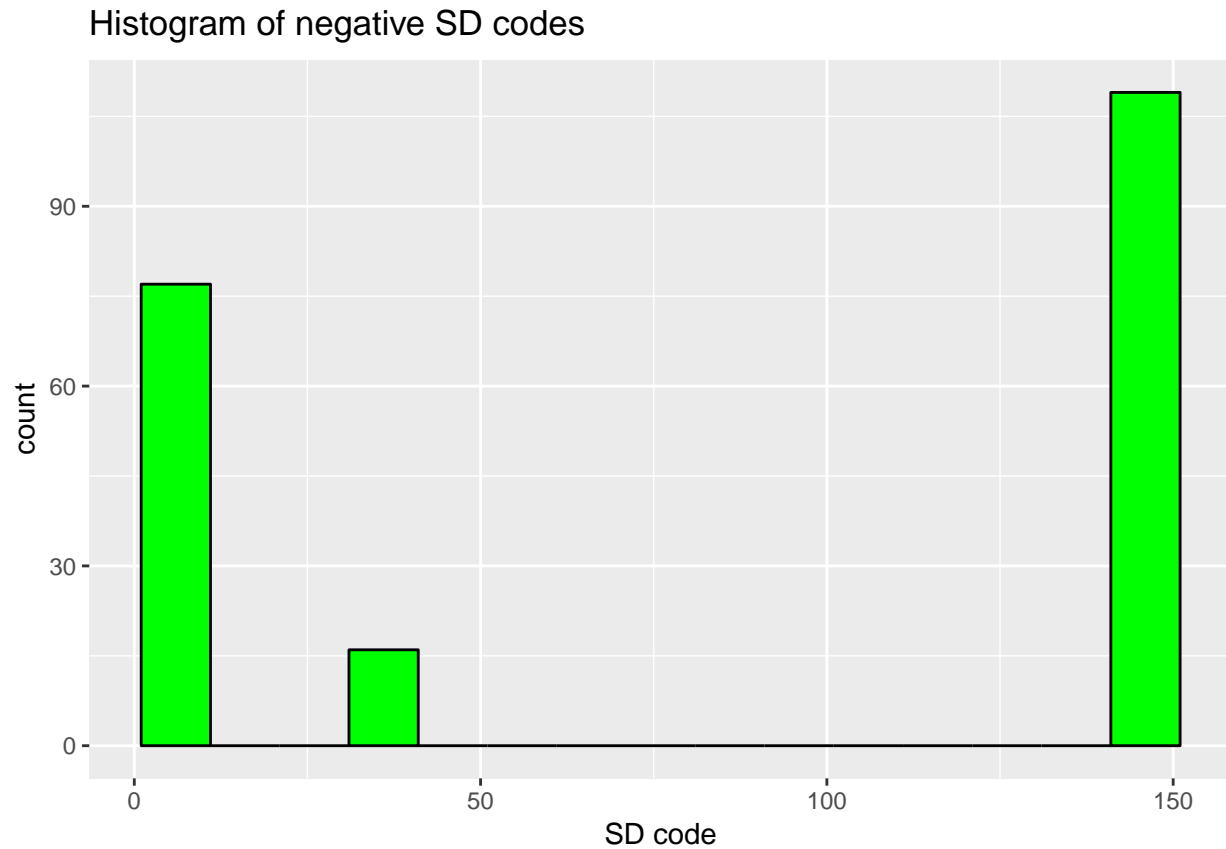
The histogram shows that many codes refer to: nature of the woodland (5 - 11), other tree species not recorded in plots (29 - 48), stumps (50 - 60). Codes 150 - 200 mainly refer to animals, the larger 190 - 200 being birds.

What about potential for codes which may have a negative association with diversity; planted conifers, rhododendron, bracken, nettle clump, (maybe road or buildings??)

```
NegativeCodesSites = c(6,41,143,145,149)
NegativeSites = SiteDescriptorsYr2 %>% select(SITE, SD_code)%>% filter(SD_code %in% NegativeCodesSites)
NegativeSites
```

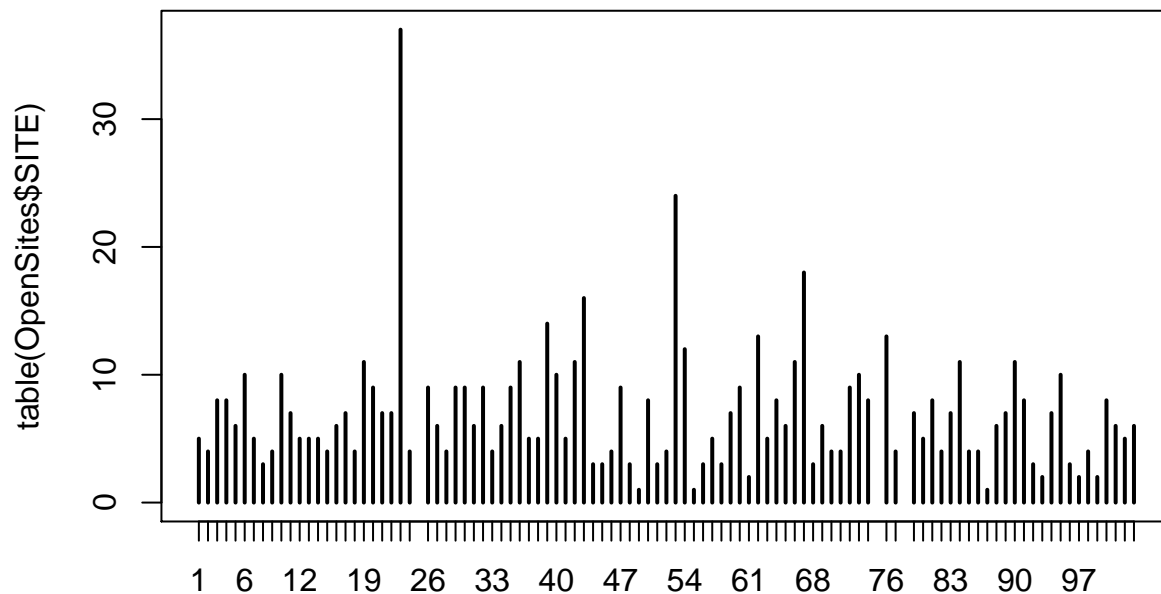
```
## # A tibble: 202 x 2
##   SITE SD_code
##   <int> <int>
## 1     3    41
## 2     1     6
## 3     1   149
## 4     2     6
## 5     2   145
## 6     3     6
## 7     3   143
## 8     3   145
## 9     3   149
## 10    4   149
## # ... with 192 more rows
```

```
ggplot(NegativeSites, aes(x = SD_code)) +
  geom_histogram(breaks = seq(1,160, by = 10), col = "black", fill = "green")+
  labs(title = "Histogram of negative SD codes", x = "SD code")
```



How many open codes does each wood have?

```
plot(table(OpenSites$SITE))
```



Most sites have around 10 SD\_codes , Site 23 has 37. Site 25 has no codes recorded, 75, 78 have no openness codes at all.

In summary: These are the SD codes which I think should be used to contribute to (or form on their own) heterogeneity index - which is linked to openness of habitat - it does not imply different habitats, such as

rotting wood, aquatics. So marsh is included because it would imply open ground.

8/9/10, coppice stool, single coppice, recently cut coppice. 24 Fire sites 86/87 Pool > 1m<sup>2</sup> (85, pool <1m<sup>2</sup> excluded) 90/91/92/93 Streams and rivers > 1m across ( streams < 1m across excluded) 100 Spring 101 Marsh 102/103 Ditches 105 - 110, glades 111/112 Rocky knolls 113 Field 114/115/116/117 paths > 1m

Of these I am not convinced about the spring and ditches. Possibly also marsh, how big is a marsh?

What might the index be - a score of the number of SD\_codes in a site?

```
ScoreSites = OpenSites %>% group_by(SITE)%>% summarise(no_rows = length(SITE))
max(ScoreSites$no_rows)
```

```
## [1] 37
```

```
mean(ScoreSites$no_rows)
```

```
## [1] 6.96
```

```
min(ScoreSites$no_rows)
```

```
## [1] 1
```

```
ScoreSites
```

```
## # A tibble: 100 x 2
```

```
##   SITE no_rows
```

```
##   <int>   <int>
```

```
## 1     1     5
```

```
## 2     2     4
```

```
## 3     3     8
```

```
## 4     4     8
```

```
## 5     5     6
```

```
## 6     6    10
```

```
## 7     7     5
```

```
## 8     8     3
```

```
## 9     9     4
```

```
## 10    10    10
```

```
## # ... with 90 more rows
```

```
#or..more concisely!!
```

```
ScoresSites = table(OpenSites$SITE)
```

The index would vary from 37 with a mean of 6.96. The vector of values Score or Scores could be added as an additional column in the Site dataframe.