Sam Norton Rbasics

Store data

tgpp <- read.csv('https://raw.githubusercontent.com/dmcglinn/quant\_methods/gh-pages/data/tgpp.csv')

Begin to view and interact

tgpp <- read.csv("tgpp.csv")  
head(tgpp)

## plot year record\_id corner scale richness easting northing slope ph  
## 1 205 1998 187 NA 100 60 727000 4080000 3 6.9  
## 2 205 1998 188 1 10 36 727000 4080000 3 6.9  
## 3 205 1998 189 2 10 34 727000 4080000 3 6.9  
## 4 205 1998 190 3 10 37 727000 4080000 3 6.9  
## 5 205 1998 191 4 10 33 727000 4080000 3 6.9  
## 6 205 1998 192 1 1 21 727000 4080000 3 6.9  
## yrsslb  
## 1 0.39  
## 2 0.39  
## 3 0.39  
## 4 0.39  
## 5 0.39  
## 6 0.39

names(tgpp)

## [1] "plot" "year" "record\_id" "corner" "scale"   
## [6] "richness" "easting" "northing" "slope" "ph"   
## [11] "yrsslb"

str(tgpp)

## 'data.frame': 4080 obs. of 11 variables:  
## $ plot : int 205 205 205 205 205 205 205 205 205 205 ...  
## $ year : int 1998 1998 1998 1998 1998 1998 1998 1998 1998 1998 ...  
## $ record\_id: int 187 188 189 190 191 192 193 194 195 196 ...  
## $ corner : int NA 1 2 3 4 1 2 3 4 1 ...  
## $ scale : num 100 10 10 10 10 1 1 1 1 0.1 ...  
## $ richness : int 60 36 34 37 33 21 23 19 25 10 ...  
## $ easting : int 727000 727000 727000 727000 727000 727000 727000 727000 727000 727000 ...  
## $ northing : int 4080000 4080000 4080000 4080000 4080000 4080000 4080000 4080000 4080000 4080000 ...  
## $ slope : int 3 3 3 3 3 3 3 3 3 3 ...  
## $ ph : num 6.9 6.9 6.9 6.9 6.9 6.9 6.9 6.9 6.9 6.9 ...  
## $ yrsslb : num 0.39 0.39 0.39 0.39 0.39 0.39 0.39 0.39 0.39 0.39 ...

1. What are the names of the columns in this dataset?

“plot” “year” “record\_id” “corner” “scale” “richness” “easting” “northing” “slope” “ph” “yrsslb”

1. How many rows and columns does this data file have?

nrow(tgpp)

## [1] 4080

ncol(tgpp)

## [1] 11

Rows = 4080  
Columns = 11

1. What kind of object is each data column? Hint: checkout the function sapply().

sapply(tgpp,class)

## plot year record\_id corner scale richness easting   
## "integer" "integer" "integer" "integer" "numeric" "integer" "integer"   
## northing slope ph yrsslb   
## "integer" "integer" "numeric" "numeric"

?sapply

1. What are the values of the the datafile for rows 1, 5, and 8 at columns 3, 7, and 10

tgpp[1,3]

## [1] 187

tgpp[1,7]

## [1] 727000

tgpp[1,10]

## [1] 6.9

tgpp[5,3]

## [1] 191

tgpp[5,7]

## [1] 727000

tgpp[5,10]

## [1] 6.9

tgpp[8,3]

## [1] 194

tgpp[8,7]

## [1] 727000

tgpp[8,10]

## [1] 6.9

1. Create a pdf of the relationship between the variables “scale” and “richness”.

require(tidyverse)

## Loading required package: tidyverse

## ── Attaching packages ──────────────────────────────────────────────────────────────────────────── tidyverse 1.2.1 ──

## ✔ ggplot2 3.2.1 ✔ purrr 0.3.2  
## ✔ tibble 2.1.3 ✔ dplyr 0.8.3  
## ✔ tidyr 0.8.3 ✔ stringr 1.4.0  
## ✔ readr 1.3.1 ✔ forcats 0.4.0

## Warning: package 'ggplot2' was built under R version 3.5.2

## Warning: package 'tibble' was built under R version 3.5.2

## Warning: package 'tidyr' was built under R version 3.5.2

## Warning: package 'purrr' was built under R version 3.5.2

## Warning: package 'dplyr' was built under R version 3.5.2

## Warning: package 'stringr' was built under R version 3.5.2

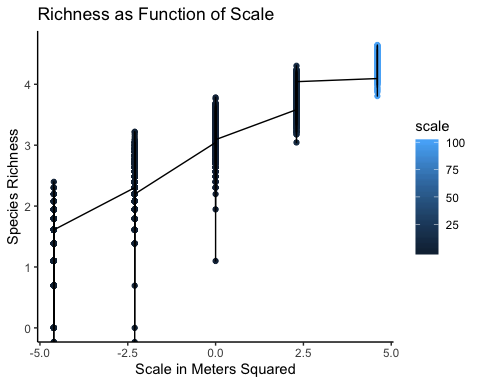
## Warning: package 'forcats' was built under R version 3.5.2

## ── Conflicts ─────────────────────────────────────────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()

names(tgpp)

## [1] "plot" "year" "record\_id" "corner" "scale"   
## [6] "richness" "easting" "northing" "slope" "ph"   
## [11] "yrsslb"

ggplot(data = tgpp) +  
 geom\_point(mapping = aes(x = log(scale), y = log(richness), color = scale)) +  
 geom\_line(mapping = aes(x = log(scale), y = log(richness))) +  
 labs(title = "Richness as Function of Scale") +  
 xlab("Scale in Meters Squared") +  
 ylab("Species Richness") +  
 theme\_classic()



ggplot(data = tgpp) +  
 geom\_boxplot(mapping = aes(x = log(scale), y = log(richness), group = scale, color = scale)) +  
 labs(title = "Species Richness as Funtion of Scale", subtitle = "Tall Grass Prarie 2010", caption = "Assignment 1 SC Norton") +  
 xlab("Scale in Meters Squared") +  
 ylab("Species Richness") +  
 theme\_minimal()

## Warning: Removed 4 rows containing non-finite values (stat\_boxplot).

