Rush Trial: data exploration 2017

Packages

library(tidyverse)

## Warning: package 'tidyverse' was built under R version 3.3.3

## Loading tidyverse: ggplot2  
## Loading tidyverse: tibble  
## Loading tidyverse: tidyr  
## Loading tidyverse: readr  
## Loading tidyverse: purrr  
## Loading tidyverse: dplyr

## Warning: package 'tibble' was built under R version 3.3.3

## Warning: package 'tidyr' was built under R version 3.3.3

## Warning: package 'readr' was built under R version 3.3.3

## Warning: package 'purrr' was built under R version 3.3.3

## Warning: package 'dplyr' was built under R version 3.3.3

## Conflicts with tidy packages ----------------------------------------------

## filter(): dplyr, stats  
## lag(): dplyr, stats

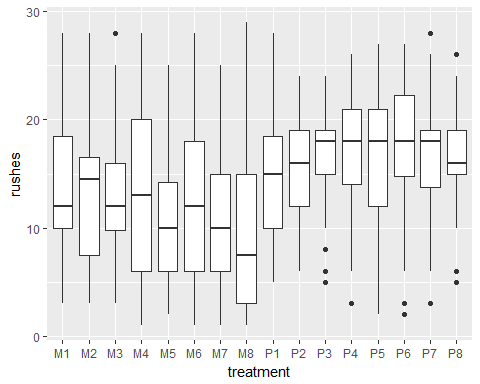
Import data

rushdata <- read.csv("../data/rushdata.csv", header = TRUE)

plotty plotty

ggplot(rushdata, aes(x = treatment, y = rushes)) +   
 geom\_boxplot(aes(colour = year))

## Warning: Removed 296 rows containing non-finite values (stat\_boxplot).

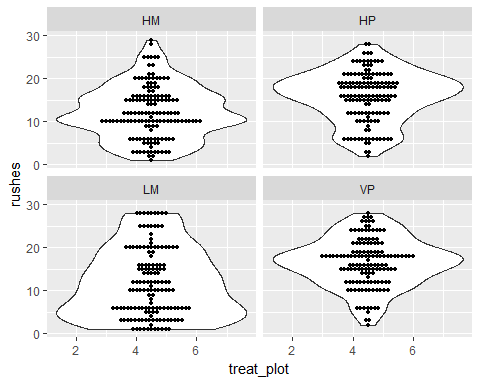


ggplot(rushdata, aes(x = treat\_plot, y = rushes)) +   
 geom\_violin(aes(colour = year)) +  
 geom\_dotplot(aes(fill = year, colour = year), binaxis='y', stackdir='center', dotsize=0.9,   
 position = position\_dodge(width = 0.9)) +  
 #geom\_point(aes(colour = year), position = position\_dodge(width = 1)) +  
 facet\_wrap(facets = ~ location)

## Warning: Removed 296 rows containing non-finite values (stat\_ydensity).

## `stat\_bindot()` using `bins = 30`. Pick better value with `binwidth`.

## Warning: Removed 296 rows containing non-finite values (stat\_bindot).

 This shows all the data for rush cover estimates, with 'violin plots' around them.

BBBut: zeroes are missing. V imp.

## Look at 2015 variability

subset 2015 data

data\_2015 <- filter(rushdata, year == 2015)

## Warning: package 'bindrcpp' was built under R version 3.3.3

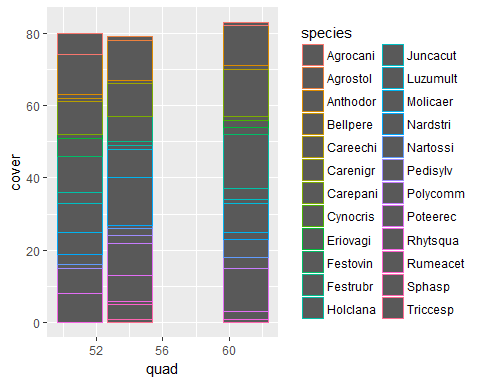
names(data\_2015)[c(1:20, ncol(data\_2015))]

## [1] "X" "location" "type" "replicate" "treatment"   
## [6] "treat\_plot" "quad" "year" "date" "surveyor"   
## [11] "rushes" "grasses" "herbs" "sedges" "bryophytes"  
## [16] "bare\_grd" "litter" "Achimill" "Agrocapi" "Agrostol"   
## [21] "Veroserp.2"

spp\_data\_2015 <- gather(data\_2015, key = "species", value = "cover", Achimill:Veroserp.2, na.rm = TRUE)  
spp\_data\_2015 <- select(spp\_data\_2015, -c(rushes:litter))

graph an example plot

ggplot(filter(spp\_data\_2015, location == "HP", replicate == "A", treatment == "P1"),   
 aes(y = cover, x = quad)) +  
 geom\_bar(aes(colour = species), stat = "identity", position = "stack")



not sure this tells me much. far too confusing. I want to show each species as a unique unit, but not necessarily identify whihc species it is.

# the categorical data   
var <- spp\_data\_2015 %>%   
 filter(location == "HP", quad == "51") %>%   
 select(species, cover)  
  
## Prep data   
nrows <- 10  
var$cover <- var$cover \* nrows^2 / 100  
  
df <- expand.grid(y = 1:nrows, x = 1:nrows)  
df$category <- factor(c(rep(var$species, var$cover),   
 rep(NA, nrows^2-length(rep(var$species, var$cover)))))   
# NOTE: if sum(var$cover) is not 100 (i.e. nrows^2), it will need adjustment to make the sum to 100.  
  
ggplot(df, aes(x = x, y = y, fill = category)) +   
 geom\_tile(color = "black", size = 0.5) +  
 scale\_x\_continuous(expand = c(0, 0)) +  
 scale\_y\_continuous(expand = c(0, 0), trans = 'reverse') +  
 labs(title="Waffle Chart", subtitle="'relative cover of species in a quadrat") +   
 scale\_fill\_discrete(na.value = "transparent")+  
 theme\_void()

