MountsBayScript.R

m281657

Thu Jul 27 14:28:07 2017

#Mounts Bay MCZ  
#attempt to implement:   
#Mount’s Bay rMCZ Intertidal Rock and Sediment Verification Survey 2013/2014  
  
#July 2017  
#Christoph Kratz  
  
setwd("C:/Christoph/Stats\_Case\_Studies/MountsBayMCZ")  
  
#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 'ggplot2' was built under R version 3.3.2

## 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

library("stringr")  
library("vegan")

## Loading required package: permute

## Loading required package: lattice

## This is vegan 2.4-1

library("ecodist")

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

## Warning: failed to assign RegisteredNativeSymbol for permute to permute  
## since permute is already defined in the 'ecodist' namespace

##   
## Attaching package: 'ecodist'

## The following object is masked from 'package:vegan':  
##   
## mantel

library("forcats")

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

#import data ----  
raw <- read.csv("data/RockQuads.csv", blank.lines.skip = T) #import csv files  
  
#remove blank lines  
raw <- raw %>%   
 filter(Quadrat != "")

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

# Check column classes  
#table(sapply(raw, class))  
#cols.num <- colnames(raw)[-c(1:7, 108, 109)]   
raw <- data.frame(lapply(raw, gsub, pattern = "%", replacement = "", fixed = T)) #remove percent dsigns  
raw <- data.frame(lapply(raw, str\_replace\_na , replacement = 0)) #change NA to 0 #change NA to 0  
  
#tidy data ----  
rock.quads <- raw %>% select(-c(X, X.1, X.2)) #drop unnecessary variables  
rock.quads[7:106] <- lapply(rock.quads[7:106], as.character) # I think this is dropping values  
rock.quads <- gather(rock.quads, 7:106, key = "species", value = "abundance", na.rm = T)#turns it into 'tidy' data  
#note that this is imperfect as you need to count columns, need to improve this  
rock.quads$abundance <- as.numeric(rock.quads$abundance) #make abundance numeric  
rock.quads <- filter(rock.quads, abundance > 0)  
  
#summarise data ----  
  
rock.quads %>%   
 group\_by(species = species) %>%   
 summarise(min = min(abundance), ave = mean(abundance), max = max(abundance), sd = sd(abundance)) %>%  
 arrange(desc(max))

## # A tibble: 57 x 5  
## species min ave max sd  
## <chr> <dbl> <dbl> <dbl> <dbl>  
## 1 Littorina.saxatilis 2 377.08333 1250 413.42702  
## 2 Spirorbis 15 463.88889 1000 370.31893  
## 3 Balanus.perforatus 2 186.00000 800 315.27068  
## 4 Chthamalus.montagui 1 45.63158 150 37.24127  
## 5 Pomatoceros.spp. 1 31.22222 100 33.84072  
## 6 Fucus.serratus 8 55.80000 96 34.17862  
## 7 Elminius.modestus 1 29.75000 90 41.03149  
## 8 Chondrus.crispus 1 23.28571 75 26.61901  
## 9 Fucus.vesiculosus 8 36.50000 75 26.20539  
## 10 Corallina.officinalis 1 19.53846 70 21.05396  
## # ... with 47 more rows

##problem here: seems to be dropping quite a few values,. Probably happens at lapply  
  
  
#row.names(rock.quads.spp) <- rock.quads.plot$Quadrat #this doesn't seem to persist in dplyr  
  
#clean data noted that quadrat data here is a mixture of percent cover and count  
#of individuals. Assume for the time being that these can be treated as  
#equivalent abundance figures, but will need to see how that goes.  
  
#subset data ----  
a1.112.quads <-   
 filter(rock.quads, Biotope == "A1.112") %>% #subset data for single biotope  
 select(-Biotope:-Comments) %>% #remove plot info  
 filter(abundance > 0) #remove blank rows  
  
#transform data  
a1.112.quads$abundance.sqrt <- sqrt(a1.112.quads$abundance)  
  
#spread data to species site matrix (for "vegan")  
a1.112.quads.spp <- spread(a1.112.quads[c("Quadrat","species", "abundance.sqrt")], key = species, value = abundance.sqrt, fill = 0)  
rownames(a1.112.quads.spp) <- a1.112.quads.spp$Quadrat  
a1.112.quads.spp$Quadrat <- NULL  
  
  
#Richness: How many species are there? ----  
#number of taxa per sample  
  
  
a1.112.quads %>% group\_by(quadrat = Quadrat) %>% count() #tabulate species number

## # A tibble: 4 x 2  
## # Groups: quadrat [4]  
## quadrat n  
## <fctr> <int>  
## 1 R10a 5  
## 2 R15b 4  
## 3 R7a 7  
## 4 R9a 3

length(unique(a1.112.quads$species)) #total number of species recorded in habitat

## [1] 11

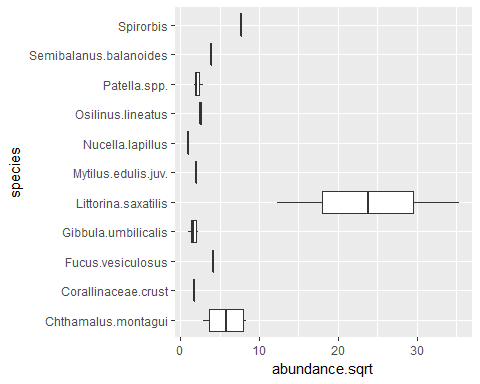
#Abundance: Which species are most or least common? ----  
#mean cover / count  
a1.112.quads %>%   
 group\_by(species = species) %>%   
 summarise(ave = mean(abundance), sd = sd(abundance)) %>%  
 arrange(desc(ave))

## # A tibble: 11 x 3  
## species ave sd  
## <chr> <dbl> <dbl>  
## 1 Littorina.saxatilis 700.00 777.8174593  
## 2 Spirorbis 60.00 NaN  
## 3 Chthamalus.montagui 38.25 31.2876440  
## 4 Fucus.vesiculosus 17.00 NaN  
## 5 Semibalanus.balanoides 15.00 NaN  
## 6 Osilinus.lineatus 6.50 0.7071068  
## 7 Patella.spp. 5.00 2.6457513  
## 8 Mytilus.edulis.juv. 4.00 NaN  
## 9 Corallinaceae.crust 3.00 NaN  
## 10 Gibbula.umbilicalis 3.00 2.8284271  
## 11 Nucella.lapillus 1.00 NaN

#mean cover / count transformed  
mean.cover.count <- a1.112.quads %>%   
 group\_by(species = species) %>%   
 summarise(ave = mean(abundance.sqrt), sd = sd(abundance.sqrt)) %>%  
 arrange(desc(ave))  
mean.cover.count

## # A tibble: 11 x 3  
## species ave sd  
## <chr> <dbl> <dbl>  
## 1 Littorina.saxatilis 23.801394 16.3397460  
## 2 Spirorbis 7.745967 NaN  
## 3 Chthamalus.montagui 5.703494 2.7616787  
## 4 Fucus.vesiculosus 4.123106 NaN  
## 5 Semibalanus.balanoides 3.872983 NaN  
## 6 Osilinus.lineatus 2.547621 0.1387779  
## 7 Patella.spp. 2.186826 0.5715665  
## 8 Mytilus.edulis.juv. 2.000000 NaN  
## 9 Corallinaceae.crust 1.732051 NaN  
## 10 Gibbula.umbilicalis 1.618034 0.8740320  
## 11 Nucella.lapillus 1.000000 NaN

levels(a1.112.quads$species) <- mean.cover.count$species[order(mean.cover.count$ave)]  
  
ggplot(data = a1.112.quads, aes(x = species, y = abundance.sqrt)) +  
 geom\_boxplot() +  
 coord\_flip()



#mean frequency  
  
# mean frequency by species and plot  
  
  
# can specify additional possible row or column levels  
  
  
# Diversity: Is the community dominated by any species, or is it even? ----  
# Shannon Wiener diversity index  
a1.112.sw <- diversity(x = a1.112.quads.spp, index = "shannon")  
  
# Margalef species richness  
  
a1.112.quads.n <-   
 a1.112.quads %>%   
 group\_by(Quadrat) %>%   
 summarise(n = length(species), N = sum(abundance), Nsqrt = sum(abundance.sqrt))  
  
a1.112.d <- (a1.112.quads.n$n -1)/log(a1.112.quads.n$N)  
   
a1.112.d <- a1.112.quads.n %>% mutate(d = (n-1)/log(N))  
  
# Pielou’s eveness  
  
#Similarity ----  
  
  
#calculate Bray Curtis dissimilarity - package 'vegan'  
a1.112.quads.bc <- vegdist(a1.112.quads.spp, method = "bray")   
a1.112.quads.bc

## R10a R15b R7a  
## R15b 0.8060240   
## R7a 0.8904280 0.4543082   
## R9a 0.4699764 0.7930035 0.8614890

summary(a1.112.quads.bc)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.4543 0.5507 0.7995 0.7125 0.8476 0.8904

#calculate Bray Curtis dissimilarity - package 'ecodist'  
a1.112.quads.bc2 <- distance(a1.112.quads.spp, method = "bray-curtis")  
a1.112.quads.bc2

## R10a R15b R7a  
## R15b 0.8060240   
## R7a 0.8904280 0.4543082   
## R9a 0.4699764 0.7930035 0.8614890

summary(a1.112.quads.bc2)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 0.4543 0.5507 0.7995 0.7125 0.8476 0.8904

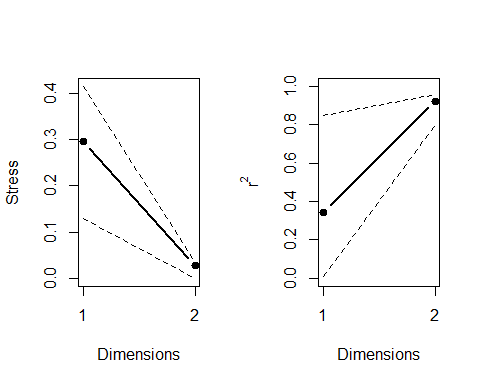
#cool, same result as using vegdist. Maybe this one is more acceptable  
#because it doesn't imply 'vegetation' in the title!!!  
  
  
#Multidimensional scaling   
a1.112.quads.nmds <- nmds(a1.112.quads.bc) #carry out NMDS on bray curtis distance  
summary(a1.112.quads.nmds)

## Length Class Mode   
## conf 20 -none- list   
## stress 20 -none- numeric  
## r2 20 -none- numeric  
## mindim 1 -none- numeric  
## maxdim 1 -none- numeric  
## nits 1 -none- numeric

a1.112.quads.nmds.min <- min(a1.112.quads.nmds, dims =2) ## choose the best two-dimensional solution to work with

## Minimum stress for given dimensionality: 8.853379e-17   
## r^2 for minimum stress configuration: 0.8011184

#this needs to be done on all teh plots, otherwise it makes little sense.   
  
  
plot(a1.112.quads.nmds)



plot(a1.112.quads.nmds.min)

