Homework_1.R

fredericronteix1

Thu Mar 8 14:44:27 2018

```
### HOMEWORK 1 ###
rm(list = ls())

setwd("~/Desktop/Taïwan/Cours/Computer statistics/1")

cop_density <- read.table("~/Desktop/Taïwan/Cours/Computer statistics/1/cop_densit
y.txt", header = T, sep = "\t")
copepod1 <- read.table("~/Desktop/Taïwan/Cours/Computer statistics/1/copepod_compo
sition.txt", header = T, sep = "\t")

rownames(cop_density) <- colnames(copepod1) #creating better names for the rows

### 1: Calculate the copepod density for each species for each cruise-station

for (i in 1:34) {
  copepod1[,i] <- (copepod1[,i]/100)*cop_density[i,1] # loop to get the density of
each species at each station
  }

copepod1 #table with the density of each species (number/m3)
```

```
##            p1      p3       p4      p6     p13      p16     p19
## 1      0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  0.0000
## 2      0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  0.0000
## 3      0.0000  0.0000   0.0000  0.0000  0.0000   4.6706 27.1206
## 4      0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  0.0000
## 5      0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  0.0000
## 6      0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  0.0000
## 7      0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  0.0000
## 8      0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  0.0000
## 9      0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  0.0000
## 10     0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  0.0000
## 11     0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  0.0000
## 12     0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  0.0000
## 13     0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  3.0134
## 14   159.5694 80.8253  40.9122 77.2920 69.2776  37.3648 21.0938
## 15     0.0000  0.0000   0.0000  0.0000  4.7348   4.6706 12.0536
## 16     0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  0.0000
## 17     0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  3.0134
## 18     0.0000  2.1907   0.0000  0.0000  0.0000   4.6706  3.0134
## 19     0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  0.0000
## 20     0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  3.0134
## 21     0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  0.0000
## 22     0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  0.0000
## 23     0.0000  0.0000   0.0000  0.0000  0.0000   0.0000  0.0000
## 160  0.000000 0.000000  0.00000 0.000000 0.000000
## 161  0.000000 0.000000  0.00000 0.000000 0.000000
## 162  0.000000 0.000000  0.13916 0.000000 0.000000
## 163  0.000000 0.000000  0.27335 0.000000 0.000000
## 164  0.000000 0.025002  0.41251 0.193050 0.583375
## 165  0.000000 0.000000  0.41251 0.386100 0.583375
## 166  0.000000 0.000000  1.23753 0.048906 0.872729
## 167  0.000000 0.000000  0.00000 0.000000 0.000000
## 168  0.000000 0.000000  0.00000 0.048906 0.000000
## 169  0.697058 0.176866  4.13007 1.256112 2.473510
## 170  0.000000 0.000000  0.00000 0.000000 0.000000
## 171  0.000000 0.000000  0.00000 0.000000 0.000000
## 172  0.000000 0.000000  0.00000 0.048906 0.000000
## 173  0.000000 0.000000  0.00000 0.000000 0.000000
## 174  0.000000 0.000000  0.00000 0.048906 0.000000
## 175  0.000000 0.000000  0.00000 0.000000 0.000000
## 176  0.000000 0.000000  0.00000 0.000000 0.000000
## 177  0.000000 0.000000  0.13916 0.000000 0.000000
## 178  0.000000 0.000000  0.00000 0.000000 0.144677
## 179  0.000000 0.000000  0.00000 0.000000 0.144677
## 180  0.000000 0.000000  0.13916 0.000000 0.000000
## 181  0.000000 0.000000  0.00000 0.000000 0.000000
```

```
## 2: For each cruise-station, calculate the species richness (number of species)
and Shannon diversity index

copepod2 <- read.table("~/Desktop/Taïwan/Cours/Computer statistics/1/copepod_compo
sition.txt", header = T, sep = "\t")
cop_density$richness <- 0 # new column with the number of species/station

# function to fill the richness column
for (j in 1:34) {
  for (i in 1:181){
    if (copepod2[i,j] != 0) {
      cop_density$richness[j] <- (cop_density$richness[j] +1) # whenever there is
a spieces in the station, the richness of the station gains +1
  }}}

cop_density$shannon <- 0 # new column for the diversity (shannon index)

cop_density$shannon <- apply(copepod2/100,2,function(x){-sum(x*log(x), na.rm = TRU
E)}) # using the apply function to calculate the shannon index

cop_density
```

```
##     density richness  shannon
## p1  1119.00        6 1.080782
## p3  1153.00       12 1.256126
## p4  1719.00        8 1.045406
## p6   855.00        9 1.114551
## p13 1246.00       31 2.144528
## p16 2123.00       29 1.413396
## p19 1159.00       43 2.582107
## p21 1497.00        7 1.566761
## p23 1351.00       46 2.994586
## p25  960.00       39 2.484830
## s18 2946.00       39 2.841237
## s19 1900.00       41 2.983345
## s20 1508.00       32 2.568974
## s22 4043.00       25 2.567104
## s23 4919.00       32 2.878602
## s25 6332.00       41 3.000568
## s27 2101.00       40 2.802674
## s29 4823.00       23 2.118423
## sA  1588.00       47 3.106234
## sB  2895.00       49 2.983319
## sC  3870.00       44 2.815962
## sD  1354.00       46 2.937975
## sE  5492.00       44 3.020831
## sF  7918.00       38 2.890280
## sG  1270.00       25 1.692271
## w22  358.10       18 1.978899
## w23  261.78       16 1.615811
## w25   80.67       24 1.841939
## w27   46.06       33 2.579666
## w29   41.74       27 2.358688
## wA     9.26       44 2.613131
## wB    49.70       54 2.999801
## wC    25.74       64 3.213878
## wD    46.67       48 3.005535
## 3: Find dominant species (species >=2% of total composition in any cruise-station) and calculate the average density for the spring, summer, and winter cruise for each dominant species.

dominant_species <- read.table("~/Desktop/Taïwan/Cours/Computer statistics/1/copepod_composition.txt", header = T, sep = "\t")
species <- data.frame((1:181)*0) # new table with only the species
names(species) <- 'dominant_species'

for (i in 1:34){
  for (j in 1:181){
    if (species [j,1] != 1){ # test to see if the species hasn't already been proved dominant
  species [j,1] <- dominant_species[j,i] >= 2 # testing the species is dominant
    }}}

species$spring_density <- 0
species$summer_density <- 0
species$winter_density <- 0 # creating new columns for the density of the dominant species

for (j in 1:nrow(species)) {
    species$spring_density [j] <- sum(copepod1[j,1:10])/10
    species$summer_density [j] <- sum(copepod1[j,11:25])/(15)
    species$winter_density [j] <- sum(copepod1[j, 26:34])/(9) # calculate the average density for the spring, summer, and winter cruise for each dominant species
  }

species <- species[-(which (species$dominant_species == 0)),]

species
```

```
##     dominant_species spring_density summer_density winter_density
## 3                  1        8.12533      26.210300     0.20313700
## 5                  1        0.00000      67.728267    23.57568900
## 14                 1       86.46599       2.500233     4.60278089
## 15                 1        3.42582     204.692020     1.26660244
## 16                 1        4.50137       1.903987     0.13661011
## 20                 1        1.27406      42.367393     0.16973356
## 35                 1       11.92657      63.691307     0.67185067
## 40                 1       11.44634      23.415760     0.68068656
## 51                 1        0.00000      28.013247     0.86376644
## 52                 1        0.00000       5.324420     1.41176278
## 54                 1        4.87340      40.484947     0.22946133
## 55                 1        0.30134       0.726680     0.42130211
## 60                 1       28.35273      41.598293    23.15445778
## 72                 1       10.71096     156.406840     0.99605678
## 73                 1        0.30720      31.915367     0.05405211
## 76                 1        5.54979       3.362153     0.00000000
## 79                 1        1.45908      26.636913     0.00000000
## 80                 1        3.74408       9.885680     0.00000000
## 81                 1        1.76042      13.302267     0.00543400
## 84                 1       54.09300      18.917167     9.39986078
## 85                 1      626.17119     414.204633    18.13617244
## 86                 1        0.00000     107.269380     0.02491267
## 88                 1       11.22283     419.718387     0.03758444
## 106                1        0.76840      20.457953     4.46338178
## 111                1        9.93754       4.661253     0.00543400
## 112                1        3.69089     267.764627     0.97009244
## 117                1        0.47348     134.511173     0.00000000
## 118                1        0.23674      53.421613     0.00000000
## 120                1        1.24800      16.452867     1.22468522
## 123                1       28.78133      83.105013     0.99051589
## 126                1      115.08926      38.426887     0.00000000
## 135                1        7.36190     197.810427     0.00000000
## 142                1      161.22854       6.445867     0.00000000
## 145                1        1.75298      97.267093     0.08584778
## 147                1        0.00000      75.731147     1.56444311
## 148                1        0.97272      21.037573     0.02334622
## 151                1        0.00000      39.107200     0.18633600
## 158                1        3.37813      16.905800     0.01086800
## 161                1       21.42939       0.000000     0.00000000
## 164                1        1.12309     116.335620     0.95004678
## 165                1        1.91489      14.383207     0.15355389
## 166                1        3.53996       4.435387     0.26321189
## 169                1       33.96802     341.826987     1.25436611
```

```r
### Notes ###

#I am not espacially good with R, I am used to using loops on python... I hope my skills will get better but for now, I am more confortable using loops than functions I do not know such as 'apply'

# CORRECTION
# spec.richness <- colSums(copepod.com>0)
# c.names <- colnames(copepod.composition)
# spring <- grep("p", c.names)
# for markdown install (knitr)
# use 'apply' function
# cop.density.e <- apply(copepod.composition, 1, function(x){cop.density*x})
# to transpose use 't(data)'
```