Hevelim Serrão de Lima

Mestranda em Oceanografia Biológica pelo Instituto Oceanográfico-IO USP

Desenvolvendo projeto com foco no fluxo de carbono na teia trófica pelágica no município de Ubatuba- SP, pelo Laboratório de Biogeoquímica Marinha-LABMAR

Laboratório de Biogeoquímica Marinha

exec

Trabalho Final

Proposta A

A identificação de organismos e a contagem destes geram grandes bancos de dados. Principalmente quando se trabalha com microorganismos, pois a contagem de indivíduos por lâmina ultrapassa 200. Administrar todas as informações sempre leva mais tempo do que o necessário. A minha proposta é tornar a análise dos dados quali- e quantitativos rápida e eficaz.

Uma função que organize mais de um banco de dados para análises qualitativas e quantitativas em uma tabela única. O objeto de entrada deve ser um arquivo csv ou txt, retornará colunas com o nome das classes,o nome das espécies, somatório total de cada espécie, média, desvio padrão. A partir disso, vai gerar um barplot apenas das classes com a porcentagem destas na amostra total.

Ex.: planilhas com nomes de espécies de microalgas, suas classes e contagem a cada amostra, poderão ser agrupadas em uma tabela única, e realizar as análises básicas de dados, somatório, média e desvio padrão.

> Oi Hevelim, preciso de mais informações pra poder te ajudar.
>
> Quais as infos que vc tem sobre cada indivíduo?
>
> A função vai ler mais de um arquivo csv/txt de uma vez e vai montar uma única tabela? E se as colunas nos bancos de dados não estiverem na mesma ordem, ou com o mesmo nome no cabeçalho?
>
> Não entendi que tipo de gráfico seria devolvido ao usuário.
>
> Não entendi que média e desvio padrão será calculado. Tem alguma coisa a ver com a abundância das espécies?
>
> Dá uma olhada nos pacotes tidyr e dplyr pra já ir tendo umas ideias, ok?
>
> bjos —-*Rena*

Proposta B

Função que calculará a taxa de predação ou herbivoria (quantitativo de presa e predador) em relação a um período de tempo determinado pelo usuário. Os objetos de entrada devem ser arquivos csv ou txt, contendo a contagem de presa e predador para cada tempo já pré-determinado pelo pesquisador (ex.: tempo 0, 6h, 12h, 18h e 24h). Será calculado o somatório total de presas e predadores para cada tempo e a partir da diferença entre estes em relação ao tempo zero, gerará um gráfico de regressão evidenciando a taxa de predação a cada período de tempo. Caso haja mais de uma espécie de presa, um barplot será criado mostrando as espécies mais consumidas pelo predador.

Ex.: Inter-relação fitoplâncton e zooplâncton A função calculará a quantidade de organismos fitoplanctônicos consumidos pelo zooplâncton dentro do ciclo circadiano, gerando um gráfico de regressão linear e um barplot com as espécies de fitoplâncton mais consumidas.

> Oi de novo,
>
> Essa pareceu mais bacanuda, hein? Ainda assim, você não conseguiu deixar a proposta muito clara.
>
> Quais devem ser os dados da entrada?
>
> O usuário precisa avisar quem é predador e quem é presa?
>
> O intervalo é um argumento da função, ou é um vetor a mais na tabela de origem?
>
> Quais serão os objetos de saída? Só o gráfico?
>
> Qual o objetivo desse tipo de análise? Estamos esgotando todas as “contas” possíveis a partir desse conjunto de dados?
>
> bjos —-*Rena*

> > Oii Rena! Desculpe o atraso na resposta. Vou tentar explicar melhor minha proposta B.
> >
> > Os dados de entrada devem ser arquivos csv ou txt com as colunas de presas e predadores separadas, o usuário precisa defini-los A ideia é ter um csv de presas e outro para predadores. Todas as espécies de presas em uma coluna e em outra coluna a sua abundância (a quantidade de cada espécie no determinado período de tempo). Portanto se a contagem de presas for feita dentro de um período de 6 em 6h em 1 dia, teremos 4 colunas (6am, 12pm, 6pm, 12am) de abundancia. O mesmo se aplica ao csv dos predadores. Pensando em predadores e presas de um ecossistema, algo mais amplo como herbívoros (zoo e fitoplâncton)
> >
> > O intervalo de tempo é mais um vetor.
> >
> > Os meus objetos de saída serão as análises estatísticas de média, desvio, variância e o gráfico mostrando a variação da abundância de presas e de predadores a cada período de tempo. E um barplot com as espécies que mais foram predadas somando os dados de todos os intervalos de tempo.

Não se ficou mais claro.

bjs

Oi Hevelim,

Entendi. Acho que pode começar a fazer essa proposta B que vai dar jogo.

Tenha carinho pelo conjunto de dados de entrada, pois talvez a forma como vc pensou em organizar te dê mais trabalho para manipular os dados.

Faça brincadeiras com essa organização antes de começar a pensar a função. vc pode colocar cada espécie como sendo uma coluna e indicar as abundâncias ao longo do tempo nas diferentes linhas

Tempo | Plecia sp. | Boa c.

12 | 20 | 2

24 | 18 | 2

36 | 14 | 3

48 | 8 | 3

Veja se entendeu minha “tabela”, pq eu não soube faze-la na wiki...

Se precisar de qq ajuda, me avise.

Beijo

Projeto Final

Função

Análise de Predação

```
setwd() # altera diretório

analise.predacao = function(file1,file2) #criando a função
{
options(digits=4) # limita número de casas decimais nos outputs
```

```
presas = read.csv(file = file1, sep=",", head=T) # lê primeiro arquivo de
dados de entrada
predadores = read.csv(file = file2, sep=",", head=T) # lê segundo arquivo de
dados de entrada

##########################################################
##### GERANDO ANÁLISE DESCRITIVA PARA PRESAS E PREDADORES
##### SOMA, MÉDIA, DESVIO-PADRAO, MÁXIMO, MEDIANA, MÍNIMO
##########################################################

#########################################
# Presas
#########################################
presas.2=presas[c(-1,-2)][presas$predacao==1,] #selecionando apenas colunas
com os períodos de tempo onde predacao=1.

soma.presas=apply(presas.2, 2, sum) # soma para cada tempo
media.presas=apply(presas.2, 2, mean) # média para cada tempo
std.presas=sqrt(apply(presas.2, 2, var)) # desvio-padrão para cada tempo
min.presas=apply(presas.2, 2, min) # mínimo para cada tempo
mediana.presas=apply(presas.2, 2, median) # mediana para cada tempo
max.presas=apply(presas.2, 2, max) # máximo para cada tempo
analise.presas=t(data.frame(soma.presas,media.presas,std.presas,min.presas,m
ediana.presas,max.presas)) # junta resultados de presas em um data.frame
rownames(analise.presas)=c("Soma","Média","Desvio-
Padrão","Mínimo","Mediana","Máximo")
print("Análise Descritiva Presas") # exibe o "título" da tabela
print(analise.presas) # exibe output no console
View(analise.presas) # exibe output em uma janela

#########################################
# Predadores
#########################################
predadores.2=predadores[c(-1)] #selecionando apenas colunas com os períodos
de tempo

soma.predadores=apply(predadores.2, 2, sum) # soma para cada tempo
media.predadores=apply(predadores.2, 2, mean) # média para cada tempo
std.predadores=sqrt(apply(predadores.2, 2, var)) # desvio-padrão para cada
tempo
min.predadores=apply(predadores.2, 2, min) # mínimo para cada tempo
mediana.predadores=apply(predadores.2, 2, median) # mediana para cada tempo
max.predadores=apply(predadores.2, 2, max) # máximo para cada tempo
analise.pred=t(data.frame(soma.predadores,media.predadores,std.predadores,mi
n.predadores,mediana.predadores,max.predadores)) # junta resultados de
predadores em um data.frame
rownames(analise.pred)=c("Soma","Média","Desvio-
Padrão","Mínimo","Mediana","Máximo")
print("Análise Descritiva predadores") # exibe o "título" da tabela
print(analise.pred) # exibe output no console
```

```
View(analise.pred) # exibe output em uma janela

########################################
### Gráfico 1 evidenciando a soma de predadores e presas observadas em cada
período de tempo
########################################

### argumentos para o "plot"
tempo=names(soma.presas) # nomes dos intervalos de tempo
ymin=min(soma.presas,soma.predadores) # valor mínimo do eixo "y" do gráfico
ymax=max(soma.presas,soma.predadores) # valor máximo do eixo "y" do gráfico
x.axis=1:length(soma.presas) # intervalos no eixo "x"

#Grafico1
png(filename="presasvspredadores.png") #Salva gráfico no diretório atual
plot(x.axis,soma.presas,
     xlab="Tempo", ylab="Quantidade de Individuos",ylim = c(ymin,ymax),
     xaxt="n", type='o', col="green", lwd=1.5) # plota gráfico
axis(1,at=1:length(tempo), labels=tempo) # nomes dos intervalos no eixo "x"
points(1:length(soma.presas),soma.predadores,pch=2, type='o', col="red",
lwd=1.5) #adiciona linha de predadores no gráfico
legend("topleft", c("Presas","Predadores"),col=c("green","red"),
pch=c(1,2),cex = 0.8) # adiciona legenda
dev.off() # retira exibição do gráfico


#Exibe gráfico no R
plot(x.axis,soma.presas,
     xlab="Tempo", ylab="Quantidade de Individuos",ylim = c(ymin,ymax),
     xaxt="n", type='o', col="green", lwd=1.5) # plota gráfico
axis(1,at=1:length(tempo), labels=tempo) # nomes dos intervalos no eixo "x"
points(1:length(soma.presas),soma.predadores,pch=2, type='o', col="red",
lwd=1.5) #adiciona linha de predadores no gráfico
legend("topleft", c("Presas","Predadores"),col=c("green","red"),
pch=c(1,2),cex = 0.8) # adiciona legenda

# Criando tabela com consumo de presas
presas=presas[with(presas,order(-predacao,generos)),] #ordena tabela em
ordem decrescente de "predacao" e depois em ordem alfabética de "generos"
presas1=presas[c(-1,-2)][presas$predacao==1,] # seleciona tempos com
"predacao"=1
presas0=presas[c(-1,-2)][presas$predacao==0,] # seleciona tempos com
"predacao"=0
diferenca=presas0-presas1 # faz a diferença das quantidades de presas em
cada tempo obtendo o consumo das presas
generos=presas[presas$predacao==1,1] # seleciona apenas generos
consumo.presas=data.frame(generos,diferenca$t0,diferenca$t1,diferenca$t2,dif
erenca$t3,diferenca$t4) # cria data.frame de consumo
colnames(consumo.presas)=names(presas[c(-2)]) # nomeia as colunas

soma.consumo.presas=data.frame(consumo.presas,rowSums(consumo.presas[c(-1)])
```

```
) # cria coluna de soma
colnames(soma.consumo.presas)=c(names(consumo.presas),"total") # nomeia
colunas
print("Consumo de Presas") # exibe "título" da tabela de consumo
print(soma.consumo.presas) # exibe tabela de consumo
View(soma.consumo.presas) # exibe tabela de em uma janela

#write.table(consumo.presas,"soma.consumo.presas.csv", sep=",")
#print(soma.consumo.presas,row.names = FALSE)

#Top 10 generos mais consumidos
maiores.consumos=head(soma.consumo.presas[with(soma.consumo.presas,order(-
total)),],10)
#Top 10 generos menos consumidos
menores.consumos=head(soma.consumo.presas[with(soma.consumo.presas,order(tot
al)),],10)
ymax=max(soma.consumo.presas$total) # valor máximo eixo "y"

#Grafico 2
png(filename="maisconsumidos.png") # salva gráfico dos mais consumidos
print("Presas mais consumidas") # "título" gráfico mais consumidos
par(las=2) # altera direção dos labels do eixo "x", posição vertical
barplot(maiores.consumos$total,names.arg=maiores.consumos$generos,col =
"blue", cex.names=0.6,ylim = c(0,ymax*1.25)) # exibe barplot dos mais
consumidos
#dev.off() # retira exibição do gráfico
dev.new()
par(las=2) # altera direção dos labels do eixo "x", posição vertical
barplot(maiores.consumos$total,names.arg=maiores.consumos$generos,col =
"blue", cex.names=0.6,ylim = c(0,ymax*1.25)) # exibe barplot dos mais
consumidos


#Grafico 3
png(filename="menosconsumidas.png") # salva gráfico dos mais consumidos
print("Presas menos consumidas")
par(las=2) # altera direção dos labels do eixo "x", posição vertical
barplot(menores.consumos$total,names.arg=menores.consumos$generos,col =
"blue", cex.names=0.6,ylim = c(0,ymax*1.25)) # exibe barplot dos menos
consumidos
#dev.off() # retira exibição do gráfico

dev.new()
par(las=2) # altera direção dos labels do eixo "x", posição vertical
barplot(menores.consumos$total,names.arg=menores.consumos$generos,col =
"blue", cex.names=0.6,ylim = c(0,ymax*1.25)) # exibe barplot dos menos
consumidos


}
```

```
#################################################################
```

Help da função

```
analise.predacao()               package:unknown               R
Documentation

            Análise descritiva, cálculo do consumo de presas e visualização
gráfica

Description:

Realiza as análises descritivas de dados de presas e predadores e calcula o
consumo das presas em função de diferentes
períodos de tempo e os exibe em tabelas. A partir da somatório individual de
presas e predadores gera um gráfico de
linhas evidenciando a quantidade de indivíduos em cada período de tempo
estabelecido. Gera dois gráficos de barras
destacando até as dez espécies de presas mais consumidas e menos consumidas
respectivamente.

Usage:

 analise.predacao("file1.csv", "file2.csv")
Arguments:

file1.csv
        Um arquivo do tipo .csv com o número de espécies de presas
observadas em cada tempo com e sem predação
File2.csv
        Um arquivo do tipo .csv com dados de predadores


Details:
O arquivo file.1 deve conter dados de um experimento que consiste de um
grupo controle, onde as presas foram observadas
em habitats sem predadores para comparar com um grupo com a presença de
predadores.
Deve conter na primeira coluna o nome dos gêneros ou espécies, na segunda
coluna deve conter a informação de
predação apenas com 1 e 0, sendo 1="com predadores no habitat" e 0="sem a
presença de predadores". As demais colunas
devem representar a quantidade de presas observadas em cada período de
tempo.


Author(s):

Hevelim Serrao de Lima
```

```
hevelimsl@usp.br
São Paulo, 25 de Junho, 2017.


Examples:

##Utilizando experimento com gêneros de fitoplâncton (presas) e zooplâncton
(predadores) seguindo o ciclo circadiano.
analise.predacao("presas.n.csv","predadores.n.csv")
```

Exemplo

presas.n.csv predadores.n.csv

From:
http://ecor.ib.usp.br/ - **ecoR**

Permanent link:
http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2017:alunos:trabalho_final:hevelimsl:start

Last update: **2020/08/12 06:04**