

Renata de Paula Orofino



Doutoranda em Ecologia, IB-USP.

Minha tese, intitulada "Argumentação de ecólogos em formação", é orientada pela Profa. Daniela Lopes Scarpa¹.

Meus exercícios

Página dos [exec](#) aqui.

Proposta de trabalho final

Página das [Propostas](#) aqui.

Código final

[Função](#)

[Help](#)

[Arquivo base para testes](#)

```
#x palavra ou termo a ser buscado
#file arquivo .txt em que se deseja realizar a busca
#exato se o termo deve ser procurado exatamente como ofi escrito
#sep separador de a ser usado na leitura do arquivo .txt indicado
#encoding tipo de encoding do arquivo.txt indicado
#ignore.acento se a busca deve ignorar acentos ou caracteres especiais
busca <- function(x, file, exato= FALSE, sep=" ", encoding="unknown",
ignore.acento= FALSE)
{
  if (sep== " ")#se o separador indicado for espaço, a função gera
apenas um produto
  {
    file1 <- scan(file, what="character", sep=sep,
encoding=encoding) #leitura do arquivo texto com encoding e separador
indicados pelo usuario#
    if (ignore.acento==TRUE) #para que a busca ignore os acentos do
texto lido e na palavra de busca
    {
```

```
letras.acento <-
c("á","à","ã","ä","â","é","è","ê","ë","í","ì","ï",
"î","ó","ò","õ","ô","ö","ú","ù","ü","û","ñ","ç","'","~")#caracteres com
acento

letras.limpo <-
c("a","a","a","a","a","e","e","e","e","i","i","i","i","o","o","o","o","o","u",
"u","u","u","n","c",""," ")#caracteres respectivos sem acento
for(l in 1:length(letras.acento)) #ciclo para substituir
acentos no texto a ser buscado #um ciclo de 1 até o tamanho das letras com
acento
{
for(v in 1:length(file1)) #um ciclo de 1 até o
tamanho do arquivo, seja ele qual for
{
temp <- grep (letras.acento[l], file1[v])
#ache o padrão com acentos no arquivo base
if (length(temp)!=0) #se houver acentos, o
tamanho será maior que 0
{
file1[v] <- gsub(letras.acento[l],
letras.limpo[l], file1[v]) #substituir os acentos no texto base por
caracteres sem acento
}
}
}
for(l in 1:length(letras.acento)) #ciclo para substituir
acentos em todo no vetor do padrão #um ciclo de 1 até o tamanho das letras
com acento
{
for(v in 1:length(x)) #um ciclo de 1 até o tamanho
do padrão de busca
{
temp <- grep(letras.acento[l], x[v]) #ache o
padrão com acentos no padrão de busca
if(length(temp)!=0) #se houver acentos, o
tamanho será maior que 0
{
x[v] <- gsub(letras.acento[l],
letras.limpo[l], x[v]) #substituir os acentos no padrão de busca por
caracteres sem acento
}
}
}
}
if (exato==FALSE) #para que a busca seja para quaisquer palavras
que contenha o padrão de busca indicado
{
posicao <- grep(pattern=x,file1,ignore.case=TRUE) #criar
um objeto com o número da linha em que o padrão foi encontrado
linha <- grep(pattern=x,file1,
```

```

value=TRUE,ignore.case=TRUE) #criar um objeto com a reprodução da linha em
que o padrão foi encontrado
      contem.palavra <- data.frame( posicao, linha) #Juntar
essas informações em uma tabela
      contagem <- paste(summary(grep(pattern=x,file1,
value=TRUE, ignore.case=TRUE))[1],"resultado(s)") # resumir quantas palavras
contém o padrão de busca indicado
      resultado1 <- list (contagem, contem.palavra) #objeto que
resumo as informações a serem devolvidas
    }
    if (exato==TRUE) #para que a busca seja por palavras exatamente
iguais à palavra indicada no padrão de busca
    {
      posicao <- grep(pattern=paste("\\b",x,"\\b",
sep=""),file1,ignore.case=TRUE) #criar um objeto com o número da linha em
que o padrão foi encontrado
      linha <- grep(pattern=paste("\\b",x,"\\b", sep=""),file1,
value=TRUE,ignore.case=TRUE) #criar um objeto com a reprodução da linha em
que o padrão foi encontrado
      palavra.exata <- data.frame( posicao, linha) #Juntar essas
informações em uma tabela
      contagem.exata <-
paste(summary(grep(pattern=paste("\\b",x,"\\b", sep=""),file1, value=TRUE,
ignore.case=TRUE))[1],"resultado(s)") # resumir quantas palavras
correspondem exatamente ao padrão de busca indicado
      resultado1 <- list(contagem.exata, palavra.exata) #objeto
que resume as informações a serem devolvidas
    }
    return (resultado1) #informações a serem retornadas ao usuário
  }
  else
  {
    file1 <- scan(file, what="character", sep=sep,
encoding=encoding) #leitura do arquivo texto com encoding e separador
indicados pelo usuario, se o separador for diferente de espaço#
    if(ignore.acento==TRUE) #para que a busca ignore os acentos do
texto lido e na palavra de busca
    {
      letras.acento <-
c("á","à","ã","ä","â","é","è","ê","ë","í","ì","ï",
"î","ó","ò","õ","ô","ö","ú","ù","ü","û","ñ","ç","'","~")#caracteres com
acento
      letras.limpo <-
c("a","a","a","a","a","e","e","e","e","i","i","i","i","o","o","o","o","o","u",
"u","u","u","n","c",""," ")#caracteres respectivos sem acento
      for(l in 1:length(letras.acento)) #ciclo para substituir
acentos no texto a ser buscado #um ciclo de 1 até o tamanho das letras com
acento
    {
      for(v in 1:length(file1)) #um ciclo de 1 até o
tamanho do arquivo, seja ele qual for

```

```
        {
            temp <- grep (letras.acento[l], file1[v])
#ache o padrão com acentos no arquivo base
            if (length(temp)!=0) #se houver acentos, o
tamanho será maior que 0
                {
                    file1[v] <- gsub(letras.acento[l],
letras.limpo[l], file1[v]) #substituir os acentos no texto base por
caracteres sem acento
                }
        }
        for(l in 1:length(letras.acento)) #ciclo para substituir
acentos em todo no vetor do padrão #um ciclo de 1 até o tamanho das letras
com acento
            {
                for(v in 1:length(x)) #um ciclo de 1 até o tamanho
do padrão de busca
                    {
                        temp <- grep(letras.acento[l], x[v]) #ache o
padrão com acentos no padrão de busca
                        if(length(temp)!=0) #se houver acentos, o
tamanho será maior que 0
                            {
                                x[v] <- gsub(letras.acento[l],
letras.limpo[l], x[v]) #substituir os acentos no padrão de busca por
caracteres sem acento
                            }
                    }
            }
        }
        if (exato==FALSE) #para que a busca seja para quaisquer palavras
que contenha o padrão de busca indicado
            {
                posicao <- grep(pattern=x,file1,ignore.case=TRUE) #criar
um objeto com o número da linha em que o padrão foi encontrado
                linha <- grep(pattern=x,file1,
value=TRUE,ignore.case=TRUE) #criar um objeto com a reprodução da linha em
que o padrão foi encontrado
                contem.palavra <- data.frame( posicao, linha) #Juntar
essas informações em uma tabela
                contagem <- paste(summary(grep(pattern=x,file1,
value=TRUE, ignore.case=TRUE))[1],"linha(s) contém palavra buscada") #
resumir quantas linhas do texto contém o padrão de busca indicado
                resultado1 <- list (contagem, contem.palavra) #objeto que
resume as informações a serem devolvidas
            }
        if (exato==TRUE) #para que a busca seja por palavras exatamente
iguais à palavra indicada no padrão de busca
            {
```

```

        posicao <- grep(pattern=paste("\\b",x,"\\b",
sep=""),file1,ignore.case=TRUE) #criar um objeto com o número da linha em
que o padrão foi encontrado
        linha <- grep(pattern=paste("\\b",x,"\\b", sep=""),file1,
value=TRUE,ignore.case=TRUE) #criar um objeto com a reprodução da linha em
que o padrão foi encontrado
        palavra.exata <- data.frame( posicao, linha) #Juntar essas
informações em uma tabela
        contagem.exata <-
paste(summary(grep(pattern=paste("\\b",x,"\\b", sep=""),file1, value=TRUE,
ignore.case=TRUE))[1],"linha(s) contém palavra buscada") # resumir quantas
linhas do texto contém a palavra correspondente ao padrão de busca indicado
        resultado1 <- list(contagem.exata, palavra.exata) #objeto
que resume as informações a serem devolvidas
    }
    file2 <- scan(file, what="character", sep=" ",
encoding=encoding) #segunda leitura do mesmo texto, para indicar os termos
isolados de sua linha de leitura
    if(ignore.acento==TRUE) #para que a busca ignore os acentos do
texto lido e na palavra de busca
    {
        letras.acento <-
c("á","à","ã","ä","â","é","è","ê","ë","í","ì","ï",
"î","ó","ò","õ","ô","ö","ú","ù","ü","û","ñ","ç","'","~")#caracteres com
acento
        letras.limpo <-
c("a","a","a","a","a","e","e","e","e","i","i","i","i","o","o","o","o","o","u",
"u","u","u","n","c",""," ")#caracteres respectivos sem acento
        for(l in 1:length(letras.acento)) #ciclo para substituir
acentos no texto a ser buscado #um ciclo de 1 até o tamanho das letras com
acento
        {
            for(v in 1:length(file2)) #um ciclo de 1 até o
tamanho do arquivo, seja ele qual for
            {
                temp <- grep (letras.acento[l], file2[v])
#ache o padrão com acentos no arquivo base
                if (length(temp)!=0) #se houver acentos, o
tamanho será maior que 0
                {
                    file2[v] <- gsub(letras.acento[l],
letras.limpo[l], file2[v]) #substituir os acentos no texto base por
caracteres sem acento
                }
            }
        }
        for(l in 1:length(letras.acento)) #ciclo para substituir
acentos em todo no vetor do padrão #um ciclo de 1 até o tamanho das letras
com acento
        {
            for(v in 1:length(x)) #um ciclo de 1 até o tamanho

```

```
do padrão de busca
    {
        temp <- grep(letras.acento[l], x[v]) #ache o
padrão com acentos no padrão de busca
        if(length(temp)!=0) #se houver acentos, o
tamanho será maior que 0
        {
            x[v] <- gsub(letras.acento[l],
letras.limpo[l], x[v]) #substituir os acentos no padrão de busca por
caracteres sem acento
        }
    }
}
if (exato==FALSE) #para que a busca seja para quaisquer palavras
que contenha o padrão de busca indicado
{
    termo <- grep(pattern=x,file2,
value=TRUE,ignore.case=TRUE) #criar um objeto com a reprodução da palavra
que contém o padrão de busca indicado
    contem.palavra2 <- data.frame( termo) #representar o
resultado em forma de tabela
    contagem2 <- paste(summary(grep(pattern=x,file2,
value=TRUE, ignore.case=TRUE))[1],"resultado(s)") # resumir quantas palavras
do texto contém o padrão de busca indicado
    resultado2 <- list( contagem2, contem.palavra2) #objeto
que resume as informações a serem devolvidas
}
if (exato==TRUE) #para que a busca seja por palavras exatamente
iguais à palavra indicada no padrão de busca
{
    termo <- grep(pattern=paste("\\b",x,"\\b", sep=""),file2,
value=TRUE,ignore.case=TRUE) #criar um objeto com a reprodução da palavra
corespondente ao padrão de busca indicado
    palavra.exata2 <- data.frame( termo) #representar o
resultado em forma de tabela
    contagem.exata2 <-
paste(summary(grep(pattern=paste("\\b",x,"\\b", sep=""),file2, value=TRUE,
ignore.case=TRUE))[1],"resultado(s)") # resumir quantas palavras do texto
correspondem ao padrão de busca indicado
    resultado2 <- list(contagem.exata2, palavra.exata2)
#objeto que resume as informações a serem devolvidas
}
return (list(resultado1,resultado2) ) #informações a serem
retornadas ao usuário
}
}
```

1)

[Veja o site do grupo](#)

From:

<http://ecor.ib.usp.br/> - **ecoR**

Permanent link:

http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2015:alunos:trabalho_final:renata.paula.silva:start 

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