

Download e manipulação de dados geográficos

Murilo S. Dias, Juan P. Quimbayo, Thiago Mendes, Mariana Bender, Augusto Flores

Contents

1	Obtendo dados e criando mapas	2
1.1	Instalando pacotes	2
1.2	Dados de peixes recifais	2
1.2.1	Onde encontrar os dados	2
1.2.2	Preparando e baixando o arquivo	3
1.2.3	Exemplo baixando dados da IUCN (CASO PRECISEM!!!)	3
1.2.4	Distribuição das espécies peixes borboletas	3
2	Riqueza de espécies na América	6
	Bibliography	13

1 Obtendo dados e criando mapas

1.1 Instalando pacotes

Para realizarmos esta prática, você precisará instalar alguns pacotes. Utilize o código abaixo para instalar os pacotes *devtools*, *maptools* e *letsR*. Pode ser que demore um pouco, mas se você tiver uma conexão internet o R vai baixar e instalar todo o necessário. Copie e cole os seguintes códigos em teu script ou terminal

```
# instalando pacotes
install.packages("devtools")
install.packages("rgdal")
install.packages("maptools")
install.packages("maps")
library(devtools)
install_github("macroecology/letsR", dependencies = T)
# install.packages('letsR')#o pacote do repositório Git é o
# mais atual
```

```
# carregando pacotes usados
library(letsR)
```

```
## Loading required package: raster
```

```
## Loading required package: sp
```

```
## Loading required package: maps
```

```
library(maptools)
```

```
## Checking rgeos availability: TRUE
```

```
library(maps)
```

```
library(rgdal)
```

```
## rgdal: version: 1.4-4, (SVN revision 833)
```

```
## Geospatial Data Abstraction Library extensions to R successfully loaded
```

```
## Loaded GDAL runtime: GDAL 2.2.2, released 2017/09/15
```

```
## Path to GDAL shared files: /usr/share/gdal/2.2
```

```
## GDAL binary built with GEOS: TRUE
```

```
## Loaded PROJ.4 runtime: Rel. 4.9.2, 08 September 2015, [PJ_VERSION: 492]
```

```
## Path to PROJ.4 shared files: (autodetected)
```

```
## Linking to sp version: 1.3-1
```

1.2 Dados de peixes recifais

1.2.1 Onde encontrar os dados

Muitos são os repositórios de dados espaciais. Dados geográficos para os organismos mais comumente acessados (e.g., *mamíferos*, *anuros*, etc) estão disponíveis no seguinte link: [*http://www.iucnredlist.org/technical-documents/spatial-data*](http://www.iucnredlist.org/technical-documents/spatial-data). Verifique as possibilidades disponíveis.

Nós utilizaremos aqui os dados de *peixes borboletas* (*Chaetodontidae*) do Novo Mundo pois são poucas as espécies e os dados são leves o suficientes para uma aula prática rápida. Se quiser testar algo mais real, tente baixar os dados de *mamíferos* (cerca de 350Mb!) em casa e utilizá-lo no lugar dos peixes Mas antes, tenha certeza que teu computador é potente o suficiente para tratar tamanho volume de dados, pois você poderá perder **algumas horas (ou até dias)** para executar os exercícios descritos abaixo.

1.2.2 Preparando e baixando o arquivo

Inicialmente, precisamos criar uma pasta no diretório de trabalho que receberá o arquivo depois de terminado o download. Para criar a pasta, fazer o download do arquivo, descomprimir e finalmente carrega-lo no R será preciso executar os seguintes códigos.

```
# criar pasta
dir.create(file.path(getwd(), "/data"))
```

1.2.3 Exemplo baixando dados da IUCN (CASO PRECISEM!!!)

Em seguida, definimos o link para download dos dados (de acordo com o endereço registrado no site da IUCN), extraímos o arquivo **zip** e carregamos o arquivo **.shp** para o R.

```
URL <- "http://bit.ly/1HSxN1K" #dados IUCN cecílias
download.file(URL, destfile = "data/GYMNOPHIONA.zip")
unzip("data/GYMNOPHIONA.zip", exdir = "data/")

# se quiser outros dados, é só mudar o link e o nome do
# arquivo URL <- 'http://bit.ly/1JTrFLL'#dadosIUCN
# Anura(~300Mb)

# download.file(URL, destfile='data/anura.zip')
# unzip('data/anura.zip', exdir = 'data/')
rm(URL) #remove o link da área de trabalho
```

Agora, verifique se os dados de cecílias estão exatamente dentro da pasta */data* que acabamos de criar. Se tudo estiver correto, você deverá ver pelo menos oito arquivos com o mesmo nome, porém extensões diferentes (*.shp*, *.shx*, *.prj*, etc). Se quiser saber mais sobre isso, você precisa procurar e ler sobre **Sistema de Informação Geográficas (SIG ou GIS em inglês)** e **Cartografia**. Muito do que é feito em *Macroecologia* é baseado em dados tipo SIG. Se puder, aprenda estas ferramentas de SIG pois este é um campo em franco desenvolvimento.

1.2.4 Distribuição das espécies peixes borboletas

Para importar os dados e ver o que tem dentro deles, faça

```
# Importante o shapefile
borb <- readOGR(dsn = "data/Chaetodontidae_NewWorld.shp")
```

```
## OGR data source with driver: ESRI Shapefile
## Source: "/home/murilo/Dropbox/Disciplina UnB_2sem2019/6_Macroecologia e Biogeografia Marin
## with 29 features
```

```
## It has 27 fields
## Integer64 fields read as strings:  id_no

# gymno <-readDGR(dsn = 'data/GYMNOPHIONA.shp')

# visualizando o conteúdo dos dados
head(borb@data)

##      id_no          binomial presence origin seasonal compiler
## 0 165627 Prognathodes aculeatus      1      1          1      IUCN
## 1   4364      Chaetodon robustus      1      1          1      IUCN
## 2 165688 Chaetodon flavirostris      1      1          1      IUCN
## 3   4363 Prognathodes obliquus       1      1          1      IUCN
## 4 165646 Prognathodes guyanensis     1      1          1      IUCN
## 5 165621      Prognathodes aya       1      1          1      IUCN
##      yrcompiled          citation source
## 0      2010 International Union for Conservation of Nature (IUCN) <NA>
## 1      2010 International Union for Conservation of Nature (IUCN) <NA>
## 2      2010 International Union for Conservation of Nature (IUCN) <NA>
## 3      2010 International Union for Conservation of Nature (IUCN) <NA>
## 4      2010 International Union for Conservation of Nature (IUCN) <NA>
## 5      2010 International Union for Conservation of Nature (IUCN) <NA>
##      dist_comm island subspecies subpop          legend tax_comm kingdom
## 0      <NA>      <NA>          <NA>      <NA> Extant (resident)      <NA> ANIMALIA
## 1      <NA>      <NA>          <NA>      <NA> Extant (resident)      <NA> ANIMALIA
## 2      <NA>      <NA>          <NA>      <NA> Extant (resident)      <NA> ANIMALIA
## 3      <NA>      <NA>          <NA>      <NA> Extant (resident)      <NA> ANIMALIA
## 4      <NA>      <NA>          <NA>      <NA> Extant (resident)      <NA> ANIMALIA
## 5      <NA>      <NA>          <NA>      <NA> Extant (resident)      <NA> ANIMALIA
##      phylum          class          order_          family          genus category
## 0 CHORDATA ACTINOPTERYGII PERCIFORMES CHAETODONTIDAE Prognathodes      LC
## 1 CHORDATA ACTINOPTERYGII PERCIFORMES CHAETODONTIDAE Chaetodon      LC
## 2 CHORDATA ACTINOPTERYGII PERCIFORMES CHAETODONTIDAE Chaetodon      LC
## 3 CHORDATA ACTINOPTERYGII PERCIFORMES CHAETODONTIDAE Prognathodes      DD
## 4 CHORDATA ACTINOPTERYGII PERCIFORMES CHAETODONTIDAE Prognathodes      LC
## 5 CHORDATA ACTINOPTERYGII PERCIFORMES CHAETODONTIDAE Prognathodes      LC
##      marine terrestrial freshwater SHAPE_Leng SHAPE_Area
## 0  True      False      False      555.8      221.5
## 1  True      False      False      191.2      70.2
## 2  True      False      False      676.2      327.2
## 3  True      False      False      5.6       2.5
## 4  True      False      False      623.9     213.9
## 5  True      False      False      283.8     74.7
```

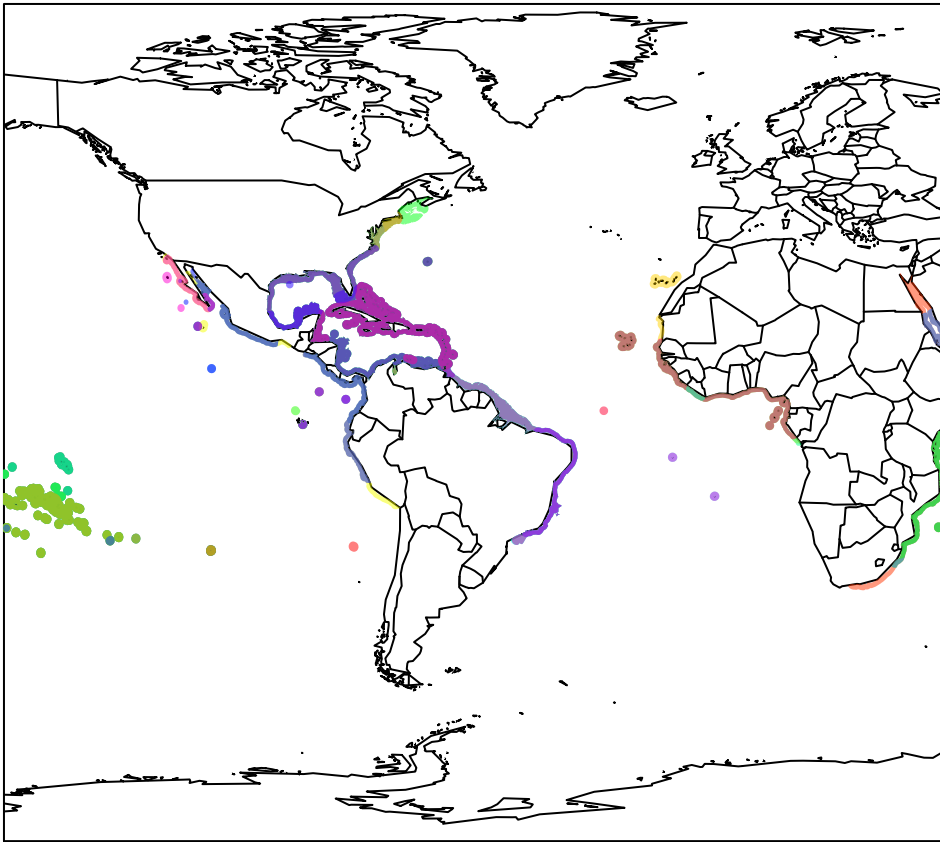
Este último comando te permitirá ver tudo todos os descritores (i.e., colunas) dentro do arquivo *.shp*.

Agora precisaremos fazer uma mapa para visualizar o que há dentro deste objeto *borb* que acabamos de criar. Utilizaremos algumas linhas de código a mais para fazer um mapa colorido para melhorar a compreensão dos dados. Utilizaremos a coluna referente ao nome científico de cada uma das espécies (i.e. *binomial*) para criar cores aleatórias e utilizá-las no nosso mapa. O parâmetro **alfa** dentro da função

rainbow faz referência ao nível de transparência de cada cor; assim, poderemos ver polígonos mesmo que eles estejam cobertos por outros polígonos.

```
## usando cores aleatórias nos nossos mapas
colors <- rainbow(length(levels(borb@data$binomial)), alpha = 0.5)
position <- match(borb@data$binomial, levels(borb@data$binomial))
colors <- colors[position]

## Plot call
map("world", xlim = c(-150, 40))
box()
plot(borb, col = colors, lty = 0, add = T)
```



Cada polígono colorido se refere à área de distribuição de cada espécie.

2 Riqueza de espécies na América

A manipulação e tratamento de dados espaciais tipo *.shp* exigem um profundo conhecimento sobre os dados e a forma como eles são armazenados. Se quiserem saber mais sobre como utilizar o R para tratar este tipo de dados, recorram a livros como **Applied spatial data analysis with R** (Bivand et al. 2013) ou **An introduction to R for spatial analysis and mapping** (Brunsdon & Comber 2015). Felizmente, existe um pacote chamado **letsR** (Vilela & Villalobos 2015) que possui uma grande variedade de ferramentas úteis para macroecólogos. Nós já instalamos e carregamos este pacote no início da aula e podemos utilizar suas funções para fazer mais alguns mapas.

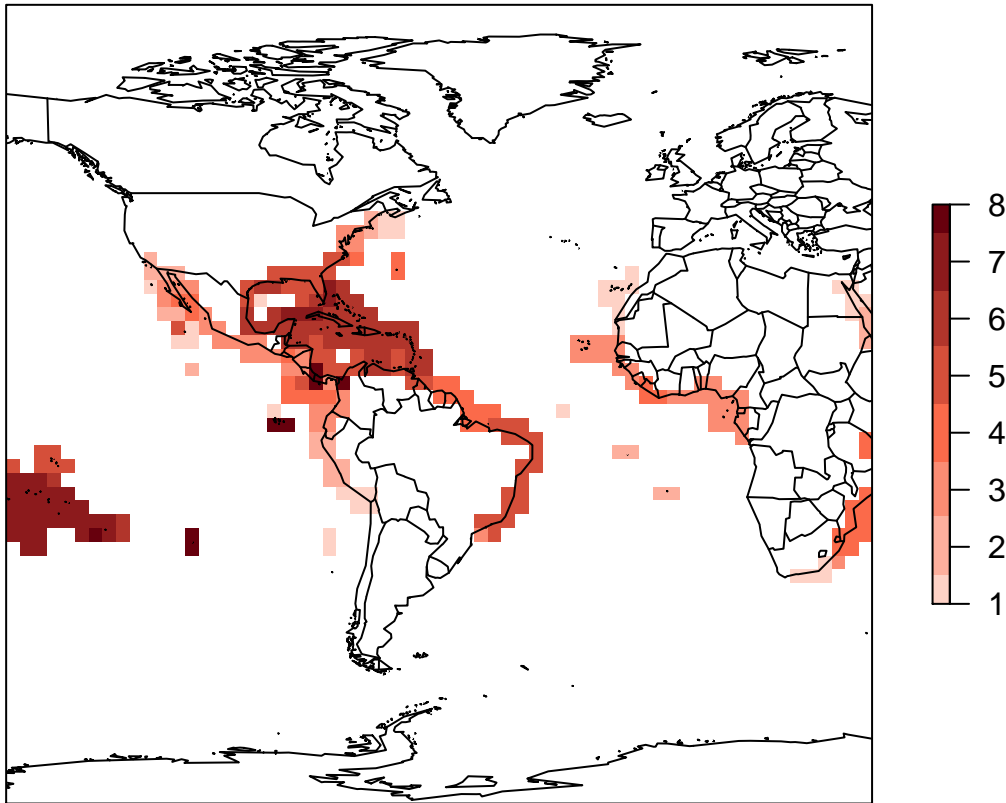
Vamos utilizar a função **lets.presab** para criar um raster (i.e., arquivo tipo imagem que armazena informações em cada um de seus pixels; digite ‘arquivo raster’ no Google para mais detalhes) resumindo a riqueza de espécie de peixes borboletas na América do Sul. Neste caso, a resolução do nosso arquivo de raster será de 3x3 graus (i.e., cada pixel ou célula tem tamanho 3 por 3 graus); mude a *resol* para valores 4 ou 5 e você verá a diferença na figura final. Já o argumento *cover* garante que somente pixels que tenham mais de 1% de sua área sobre do continente seja mantido (mude estes valores e veja a diferença no mapa resultante). Como nosso objetivo é uma mapa da América, as coordenadas limites foram definidas para realizar o exercício somente ao longo da extensão desejada (argumentos *xmn*, *xmx*, *ymn* e *ymx* da função **lets.presab**).

```
borb_maps <- lets.presab(borb, resol = 3,  
                        xmn = -150, xmx = 40, #America range  
                        ymn = -38, ymx = 45, #America range  
                        cover = 0.01)
```

```
## Warning in `proj4string<-`(`*tmp*`, value = <S4 object of class structure("CRS", package =  
## +proj=longlat +datum=WGS84 +no_defs +ellps=WGS84 +towgs84=0,0,0  
## without reprojecting.  
## For reprojection, use function spTransform
```

```
plot(borb_maps, axes=F, main="Riqueza de Peixes Borboletas")
```

Riqueza de Peixes Borboletas



Aproveite e use o seguinte código para ver o que há no objeto criado pela função

```
# see what we have
```

```
summary(borb_maps)
```

```
##  
## Class: PresenceAbsence  
## _ _  
## Number of species: 29  
## Number of cells: 289  
## Cells with presence: 289  
## Cells without presence: 0  
## Species without presence: 0  
## Species with the largest range: Chaetodon ocellatus  
## _ _  
## Grid parameters  
## Resolution: 3, 3 (x, y)  
## Extention: -150, 39, -39, 45 (xmin, xmax, ymin, ymax)  
## Coord. Ref.: +proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0
```

```
str(borb_maps)
```

```
## List of 3  
## $ Presence_and_Absence_Matrix: num [1:289, 1:31] -70.5 -67.5 -64.5 -76.5 -73.5 -70.5 -67.  
## ..- attr(*, "dimnames")=List of 2  
## .. ..$ : NULL
```

```

## .. ..$ : chr [1:31] "Longitude(x)" "Latitude(y)" "Amphichaetodon melbae" "Chaetodon auri
## $ Richness_Raster :Formal class 'RasterLayer' [package "raster"] with 12 slots
## .. ..@ file :Formal class '.RasterFile' [package "raster"] with 13 slots
## .. .. .. ..@ name : chr ""
## .. .. .. ..@ datanotation: chr "FLT4S"
## .. .. .. ..@ byteorder : chr "little"
## .. .. .. ..@ nodatavalue : num -Inf
## .. .. .. ..@ NChanged : logi FALSE
## .. .. .. ..@ nbands : int 1
## .. .. .. ..@ bandorder : chr "BIL"
## .. .. .. ..@ offset : int 0
## .. .. .. ..@ toptobottom : logi TRUE
## .. .. .. ..@ blockrows : int 0
## .. .. .. ..@ blockcols : int 0
## .. .. .. ..@ driver : chr ""
## .. .. .. ..@ open : logi FALSE
## .. ..@ data :Formal class '.SingleLayerData' [package "raster"] with 13 slots
## .. .. .. ..@ values : num [1:1764] 0 0 0 0 0 0 0 0 0 0 ...
## .. .. .. ..@ offset : num 0
## .. .. .. ..@ gain : num 1
## .. .. .. ..@ inmemory : logi TRUE
## .. .. .. ..@ fromdisk : logi FALSE
## .. .. .. ..@ isfactor : logi FALSE
## .. .. .. ..@ attributes: list()
## .. .. .. ..@ haveminmax: logi TRUE
## .. .. .. ..@ min : num 0
## .. .. .. ..@ max : num 8
## .. .. .. ..@ band : int 1
## .. .. .. ..@ unit : chr ""
## .. .. .. ..@ names : chr ""
## .. ..@ legend :Formal class '.RasterLegend' [package "raster"] with 5 slots
## .. .. .. ..@ type : chr(0)
## .. .. .. ..@ values : logi(0)
## .. .. .. ..@ color : logi(0)
## .. .. .. ..@ names : logi(0)
## .. .. .. ..@ colortable: logi(0)
## .. ..@ title : chr(0)
## .. ..@ extent :Formal class 'Extent' [package "raster"] with 4 slots
## .. .. .. ..@ xmin: num -150
## .. .. .. ..@ xmax: num 39
## .. .. .. ..@ ymin: num -39
## .. .. .. ..@ ymax: num 45
## .. ..@ rotated : logi FALSE
## .. ..@ rotation:Formal class '.Rotation' [package "raster"] with 2 slots
## .. .. .. ..@ geotrans: num(0)
## .. .. .. ..@ transfun:function ()
## .. ..@ ncols : int 63
## .. ..@ nrows : int 28
## .. ..@ crs :Formal class 'CRS' [package "sp"] with 1 slot

```



```
## .. .. .@ projargs: chr "+proj=longlat +datum=WGS84 +ellps=WGS84 +towgs84=0,0,0"
## .. ..@ history : list()
## .. ..@ z : list()
## $ Species_name : chr [1:29] "Amphichaetodon melbae" "Chaetodon auriga" "Cha
## - attr(*, "class")= chr "PresenceAbsence"
```

E se precisarmos criar somente uma matrix para saber qual espécie ocorre em cada uma das células do raster, você pode acessar a matriz direto do objeto **borb_maps** ou refazer a mesma função pedindo como resultado somente a matrix de presença/ausência (*show.matrix=TRUE* dentro da função *lets.presab*; digite *?lets.presab* no console do R e leia o help desta função e seus argumentos).

```
# Presence/absence matrix
borb_pa <- borb_maps$Presence_and_Absence_Matrix

dim(borb_pa)

## [1] 289 31

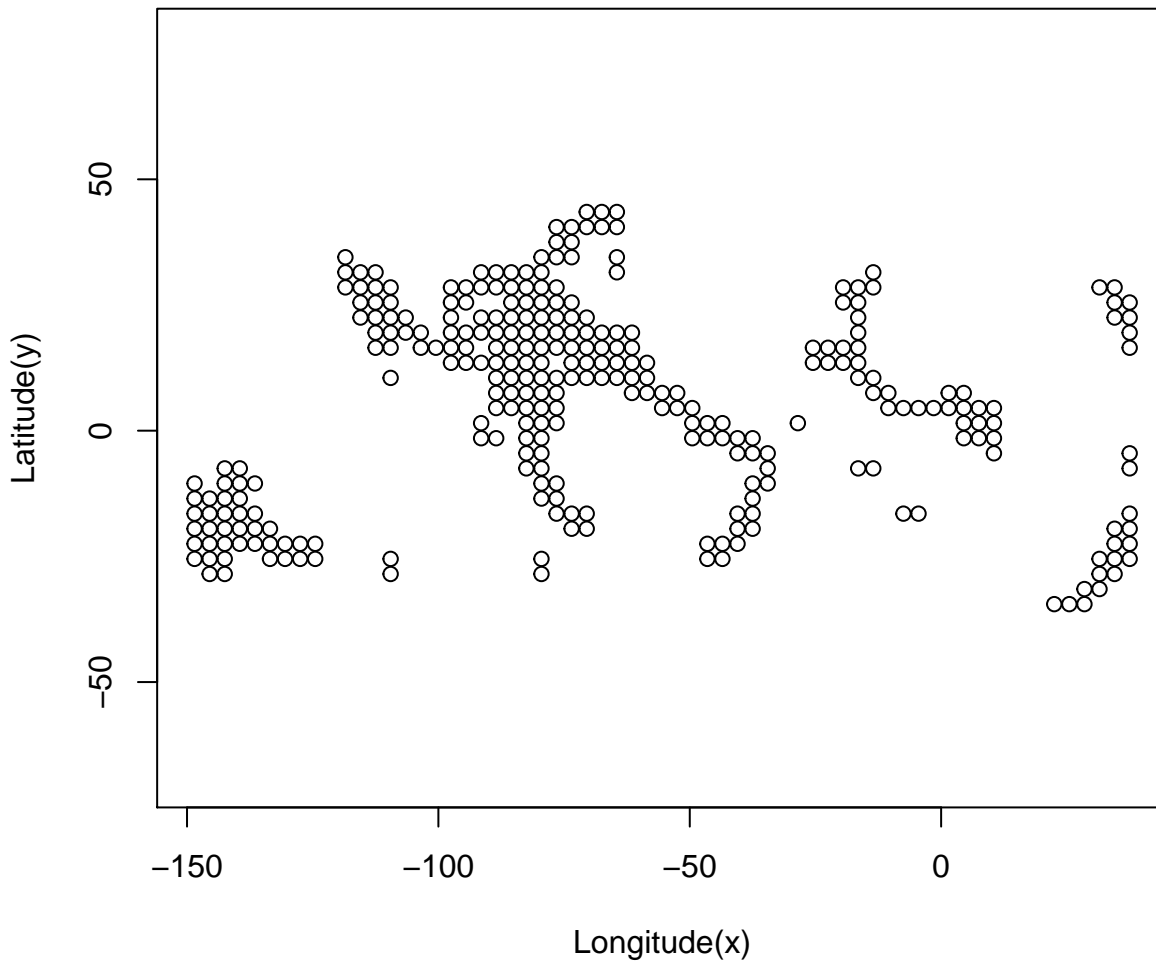
# species names
colnames(borb_pa)
```

```
## [1] "Longitude(x)" "Latitude(y)"
## [3] "Amphichaetodon melbae" "Chaetodon auriga"
## [5] "Chaetodon capistratus" "Chaetodon flavirostris"
## [7] "Chaetodon hoefleri" "Chaetodon humeralis"
## [9] "Chaetodon litus" "Chaetodon lunula"
## [11] "Chaetodon mertensii" "Chaetodon meyeri"
## [13] "Chaetodon ocellatus" "Chaetodon pelewensis"
## [15] "Chaetodon robustus" "Chaetodon sanctaehelenae"
## [17] "Chaetodon sedentarius" "Chaetodon striatus"
## [19] "Chaetodon unimaculatus" "Forcipiger flavissimus"
## [21] "Hemitaurichthys multispinosus" "Johnrandallia nigrirostris"
## [23] "Prognathodes aculeatus" "Prognathodes aya"
## [25] "Prognathodes brasiliensis" "Prognathodes carlhubbsi"
## [27] "Prognathodes dichrous" "Prognathodes falcifer"
## [29] "Prognathodes guyanensis" "Prognathodes marcellae"
## [31] "Prognathodes obliquus"
```

```
# presence/absence matrix
borb_pa[1:5, 1:5]
```

```
##      Longitude(x) Latitude(y) Amphichaetodon melbae Chaetodon auriga
## [1,]          -70          44                0                0
## [2,]          -68          44                0                0
## [3,]          -64          44                0                0
## [4,]          -76          40                0                0
## [5,]          -74          40                0                0
##      Chaetodon capistratus
## [1,]                1
## [2,]                0
## [3,]                0
```

```
## [4,] 1
## [5,] 1
# Position of cells
plot(borb_pa[, 1:2], asp = 1)
```



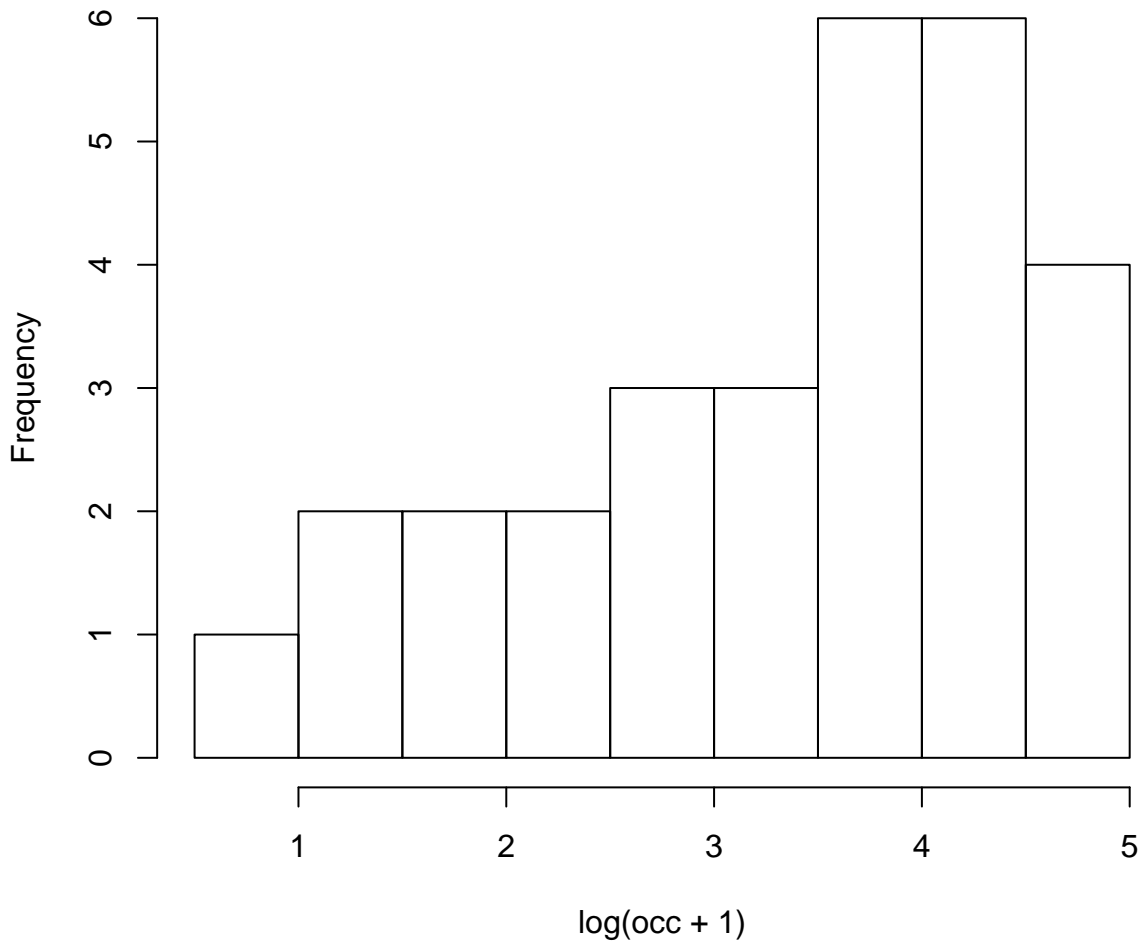
```
# species occurrence
n <- dim(borb_pa)[2]
occ <- colSums(borb_pa[, c(3:n)])
occ
```

##	Amphichaetodon melbae	Chaetodon auriga
##	4	66
##	Chaetodon capistratus	Chaetodon flavirostris
##	81	34
##	Chaetodon hoeferi	Chaetodon humeralis
##	37	61
##	Chaetodon litus	Chaetodon lunula
##	2	58
##	Chaetodon mertensii	Chaetodon meyeri
##	33	16
##	Chaetodon ocellatus	Chaetodon pelewensis
##	115	41
##	Chaetodon robustus	Chaetodon sanctaehelenae
##	29	7

##	Chaetodon sedentarius	Chaetodon striatus
##	106	108
##	Chaetodon unimaculatus	Forcipiger flavissimus
##	43	93
##	Hemitaurichthys multispinosus	Johnrandallia nigrirostris
##	3	51
##	Prognathodes aculeatus	Prognathodes aya
##	64	27
##	Prognathodes brasiliensis	Prognathodes carlhubbsi
##	18	8
##	Prognathodes dichrous	Prognathodes falcifer
##	4	12
##	Prognathodes guyanensis	Prognathodes marcellae
##	70	29
##	Prognathodes obliquus	
##	1	

```
hist(log(occ + 1), main = "Ocorrência de espécies por pixel (log)")
```

Ocorrência de espécies por pixel (log)



Se quisermos calcular a área de distribuição de cada uma das espécies, é só utilizar a seguinte função. Porém, como ela retorna o valor em m^2 , precisamos converter em km^2 dividindo por 1000000 (1000m x 1000m).

```

# range size in m2
range_gym <- lets.rangesize(borb_maps, units = "squaremeter")

# transformar m2 para km2
range_gym <- range_gym/1e+06
range_gym

```

```

##                               Range_size
## Amphichaetodon melbae         4.0e+05
## Chaetodon auriga              6.8e+06
## Chaetodon capistratus        8.2e+06
## Chaetodon flavirostris       3.5e+06
## Chaetodon hoefleri           4.0e+06
## Chaetodon humeralis          6.5e+06
## Chaetodon litus              2.0e+05
## Chaetodon lunula             6.1e+06
## Chaetodon mertensii         3.4e+06
## Chaetodon meyeri            1.7e+06
## Chaetodon ocellatus          1.2e+07
## Chaetodon pelewensis        4.3e+06
## Chaetodon robustus          3.2e+06
## Chaetodon sanctaehelenae     7.6e+05
## Chaetodon sedentarius        1.1e+07
## Chaetodon striatus           1.1e+07
## Chaetodon unimaculatus       4.5e+06
## Forcipiger flavissimus       9.8e+06
## Hemitaurichthys multispinosus 3.0e+05
## Johnrandallia nigrirostris   5.5e+06
## Prognathodes aculeatus       6.7e+06
## Prognathodes aya             2.7e+06
## Prognathodes brasiliensis    1.9e+06
## Prognathodes carlhubbsi      8.8e+05
## Prognathodes dichrous        4.3e+05
## Prognathodes falcifer        1.2e+06
## Prognathodes guyanensis      7.4e+06
## Prognathodes marcellae       3.2e+06
## Prognathodes obliquus        1.1e+05

```

Na próxima aula nós trataremos de área de distribuição dos organismos.

Bibliography

1. Vilela, B. & Villalobos, F. (2015). letsR: A new r package for data handling and analysis in macroecology. *Methods in Ecology and Evolution*, 6, 1229–1234.