



Universidade Estadual do Rio Grande do Sul

Fundação Zoobotânica do Estado do Rio Grande do Sul e  
Universidade do Estado do Rio Grande do Sul  
Mestrado em Sistemática e Conservação da biodiversidade

Introdução ao Ambiente R

Atividade Final

Prof. Dr. Márlon de Castro Vasconcelos\*

26 Março 2022

### Resumo

A atividade final da disciplina consistirá em escrever e executar alguns códigos visto em sala de aula. A atividade deverá ser reportada em arquivo criado em RMarkdown. Uma vez finalizado o código e a escrita em *RMarkdown*, o discente deverá enviar para o meu e-mail o arquivo *.rmd*.

## Conteúdo

1. Do Conjunto de Dados	2
2. Instruções Gerais	4
3. Atividade	5

---

\*marlon-vasconcelos@uergs.edu.br, <https://github.com/Vascomc>

# 1. Do Conjunto de Dados

Os dados foram obtidos a partir do **Portal da Biodiversidade**. Os parâmetros de busca foram como seguem na figura 1.

```
knitr::include_graphics("Parametros.png")
```

Nome Científico

Nome Comum

Estado de Conservação

Categoria de Ameaça

Base de Dados: sistema de autorização e informação em biodiversidade.

Período: Data de Início 26/03/2022, Data de Fim 26/03/2022

Estado/Território: rs

Jurisdição: uc federal

Responsável

Localidade

Pesquisar Limpar

Figura 1: Parâmetros de busca

Após, o download dos dados, foi realizado uma seleção de variáveis para a inclusão nessa atividade. Conforme código abaixo.

```
# Nesse chunk usei a opção {r, eval = FALSE}, assim o código é mostrado, mas  
# não executado.  
  
# Dados no mesmo diretório/Projeto  
dados <- read_csv("portalbio_export_25-03-2022-10-43-34.csv")  
  
# Selecionei as seguintes variáveis para a Atividade  
df <- dados %>%  
  select(Nome.da.instituicao, Classe, Nome.cientifico, Numero.de.individuos, Categoria.d  
    Localidade, Municipio, Latitude, Longitude)  
  
# OBS.: Lembrando que o pipe %>%, para a versão do R => 4.1 pode usar o |>  
# deste que configurado em tools > global options > code > use native pipe  
# operator, |> (Require R > 4.1 +)
```

Para o download dos dados no repositório do GitHub basta copiar e colar no **scrip** o código abaixo

```
dados <- read.csv2("https://raw.githubusercontent.com/Vascomc/introducao-ao-R-FZB-Uergs  
/babe17b83f240734362f318d5f41f94615581db2  
/Atividade_Disicplina/Atividade.csv")
```

Nesse conjunto de dados temos grupos diversos, desde Aves, passando por Tubellaria a Zygnematomyceae. Ou seja, bem diversificado. A localização são UC's ou locais nos limites delas. Eu dei uma corrigida nos nomes de modo a facilitar a vocês. Por que isso? Bom como as entradas são feitas por pessoas diferentes há situações como por exemplo: Parque... do Jacui, e Parque .... Delta do Jacuí. Temos a mesma UC, mas escrita de formas diferentes. Logo, quando baixamos um conjunto de dados de uma base que não tem uma curadoria aprofundada, ou uma formatação prévia, sempre temos que dar uma olhada nesses detalhes. **FICA A DICA.**

## 2. Instruções Gerais

---

Sugestões, para a criação dos códigos:

- Criar e rodar os códigos em um *scrip*;
- Ler com atenção o que está sendo pedido;
- Ver exemplos nas aulas 1 e 2;
- Colocar os comentários para que possa entender o que está fazendo.
- Para todos os gráficos acrescente os títulos dos eixos

Sugestões, para a criação arquivo .Rmd:

- Para o YAML
  - Título: Nome da disciplina;
  - Sub-título: Qula a atividade;
  - Autor: Nome do aluno;
  - data;
  - Saída: deverá ser em HTML
- Cada título deverá ser um exercício;
- Inserir os *chunks* e cópiar o código do *script* e colar no *chunk*
  - Atente-se para as configurações de cada *chunk*;
- No corpo do texto abaixo do código do exercício, descrever o que ele faz;
- Usar o pacote **kableExtra** para as tabelas;
- Todos os códigos deverão ser mostrados.

### 3. Atividade

---

1. Carregue os pacotes, **tidyverse**, **rstatix**, **kableExtra** e **ggpubr**;
2. Crie um objeto e transforme as variáveis *character* em *factor*, exceto Latitude e Longitude;
3. Crie uma tabela de resumo estatístico para o número de indivíduos para cada UC;
4. Crie um gráfico de barras:
  1. tendo no eixo y as UC's e no eixo x a riqueza de espécie, salve em um objeto. Aqui será necessário usar a função *summarise* como segue *summarise(riqueza = n())* <sup>1</sup>;
  2. tendo no eixo y as UC's e no eixo x a abundância de espécie, salve em um objeto;
  3. Coloque os dois gráficos juntos usando a função *ggarrange ()*.
  4. 4.1. a 4.3. no mesmo *chunk*.
  5. Quais as três UC's com maior Diversidade e Abundância? São as mesmas?
5. Crie um objeto e subdivida o conjunto de dados selecionando as três UCS's com maior abundância.
6. Passe o objeto criado em 5 para o formato *wide* e salve em um objeto, e mostre as 6 primeira linhas desse objeto como uma tabela.
7. Crie uma matriz, contendo 15 linhas e 5 colunas. Preencha as colunas usando distribuição Normal para as 3 primeiras colunas e distribuição de Poisson para as duas ultimas.
  1. se atente para os parâmetros das funções *rnorn* e *rpois*
8. Quantas espécies são **Criticamente ameaçadas**, e para quantas **não há informações**?
9. Qual a Classe com mais espécies? ela também é a mais abundante? faça o mesmo para Classe que o feito para UC's no exercício 4.
10. Baixe o loga da FZB e o insira no arquivo, sendo centralizado e com 40% de seu tamanho original.
11. Que comando usamos para que o R entenda as colunas de um conjunto de dados como se fossem objetos? exemplifique num *chunk*.
12. Baixe os dados do artigo Competition drives the response of soil microbial diversity to increased grazing by vertebrate herbivores

```
dados_2 <- read.csv2("https://github.com/Vascomc
                    /introducao-ao-R-FZB-Uergs/raw
                    /main/Atividade_Disicplina/Dryad_Eldridge_Ecology.csv")

dados_2 <- dados_2 |>
  mutate(pH = as.numeric(pH), Carbon = as.numeric(Carbon))
```

1. Crie um grafico de disperssão entre o pH e o *Carbon* e use a variável *Community* para colorir os pontos no gráfico.
2. Qual o padrão que emerge da figura resultante?
3. Qual a comunidade com menor capacidade de retenção de Carbono?
4. Use o argumento *add = "reg.line"* e apenas por inspesão visual, há regressão de pH sobre o Carbono em alguma comunidade avaliada? Que tendências você observa?

---

<sup>1</sup>Parte final do código.

13. Dados os dados abaixo:

```
ambiente <- data.frame(list(Locais = c(1:20), OD = rnorm(20, 6, 0.5), temp = rnorm(20, 23, 1), pH = rnorm(20, 7, 0.2), Cond = rnorm(20, 50, 30), estagio = gl(4, 5, labels = c("Climax", "Inicial", "Medio_tardio", "Medio_inicial"))))

comunidade <- data.frame(list(Locais = c(1:20), Sp1 = rpois(20, 2), Sp2 = rpois(20, 5), Sp3 = rpois(20, 10), Sp4 = rpois(20, 13), Sp5 = rpois(20, 13), Sp6 = rpois(20, 1), Sp7 = rpois(20, 0.5), Sp8 = rpois(20, 4)))
```

1. Como podemos junta dos dois conjuntos de dados e exeportá-los como arquivo *.csv*?
14. Usando o pacote **GGplot** crie um gráfico de erro, usando como dados, a abundância de espécies por estágio.