

# Plano de ensino

Maurício Vancine

17/11/2021

## Introdução ao R para análise de dados

### Docentes responsáveis

Prof. Maurício Humberto Vancine

### Carga horária

18 h

### Participantes

20 alunos

### Datas e horários

17/10/2021 a 19/10/2021 (9-17h)

### Repositório de dados do curso

<https://github.com/mauriciovancine/workshop-r-data-analysis>

### Resumo

O curso oferecerá os principais conceitos teóricos e práticos sobre o uso da linguagem R para manipulação, visualização e análise de dados tabulares, com enfoque em Ecologia Geral. Serão abordados os seguintes temas: (1) introdução à linguagem R, (2) introdução ao tidyverse, (3) visualização de dados, (4) introdução à estatística, (5) análise exploratória de dados, (6) modelos lineares e (7) modelos lineares generalizados e seleção de modelos. Após a realização do curso, espera-se que as alunas e alunos adquiram conceitos gerais sobre a estrutura, manipulação, visualização e análise de dados tabulares, assim como domínio das técnicas e métodos para alcançar autonomia e produzir soluções para suas próprias questões relativas à aplicação da estatística para Ecologia.

## Conteúdo

### 1 Introdução à linguagem R

1. Linguagem R
2. RStudio
3. Console
4. Scripts
5. Operadores
6. Objetos
7. Funções
8. Pacotes
9. Ajuda
10. Ambiente
11. Citações
12. Principais erros
13. Principal material de estudo
14. Atributos dos objetos
15. Manipulação de dados
16. Valores faltantes e especiais
17. Diretório de trabalho
18. Importar dados
19. Conferência de dados importados
20. Exportar dados

### 2 Introdução ao tidyverse

1. Contextualização
2. tidyverse
3. here
4. readr, readxl e writexl
5. tibble
6. magrittr (pipe - %>%)
7. tidyr
8. dplyr
9. stringr
10. forcats
11. lubridate
12. purrr

### 3 Visualização de dados

1. Contextualização
2. Pacotes para produção de gráficos
3. Gramática dos gráficos
4. Principal material de estudo
5. Principais tipos de gráficos
6. Histograma e Densidade
7. Gráfico de setores
8. Gráfico de barras
9. Gráfico de caixas
10. Gráfico de dispersão
11. Gráfico pareado

12. Combinando gráficos
13. Gráficos animados
14. Gráficos interativos
15. Gráficos usando interface

#### **4 Introdução à estatística**

1. Formulação de perguntas em Ecologia
2. Abordagens filosóficas e estatísticas em Ciência
3. Hipóteses científicas
4. Hipóteses estatísticas
5. Teste de hipóteses
6. Princípio da parcimônia
7. Delineamento amostral
8. Operacionalização em medidas

#### **5 Análise exploratória de dados**

1. Organização de dados em planilhas
2. Estatística descritiva
3. Variáveis estatísticas
4. Tipos de variáveis estatísticas
5. Distribuição de frequência de variáveis
6. Dados discrepantes (*outliers*)
7. Transformações em variáveis estatísticas

#### **6 Modelos lineares**

1. Variáveis resposta e explanatórias
2. Tipos de análises de dados
3. Análises univariadas
4. Teste-t
5. Teste-t não-paramétrico
6. Análise de variância (ANOVA)
7. Análise de variância (ANOVA) não-paramétrica
8. Modelos lineares simples
9. Modelos lineares múltiplos

#### **7 Modelos lineares generalizados e seleção de modelos**

1. Verosimilhança
2. Modelos Lineares Generalizados (GLM)
3. Principais famílias de distribuição
4. Múltiplas hipóteses concorrentes
5. Critério de Informação de Akaike (AIC)

#### **Referências**

Adler J. 2012. R in a Nutshell: A Desktop Quick Reference. 2 ed. O'Reilly Media.

- Beckerman, A., Childs, D., & Petchey, O. (2017). *Getting Started with R: An Introduction for Biologists*. Oxford University Press.
- Bolker, B. M. (2008). *Ecological models and data in R*. Princeton University Press.
- Braun, J., & Murdoch, D. J. (2016). *A first course in statistical programming with R* (Second edition). Cambridge University Press.
- Chang W. 2013. *R Graphics Cookbook: Practical Recipes for Visualizing Data*. 2 ed. O'Reilly Media. [<https://r-graphics.org/>]
- Cotton R. 2013. *Learning R: A Step-by-Step Function Guide to Data Analysis*. O'Reilly Media.
- Davies TM. 2016. *The Book of R: A First Course in Programming and Statistics*. No Starch Press.
- Engel C. 2019. *Introduction to R*. [<https://cengel.github.io/R-intro/>]
- Fox, G. A., Negrete-Yankelevich, S., & Sosa, V. J. (Orgs.). (2015). *Ecological Statistics: Contemporary theory and application*. Oxford University Press.
- Gardener, M. (2017). *Statistics for ecologists using R and Excel: Data collection, exploration, analysis and presentation* (Second edition). Pelagic Publishing.
- Gotelli, N. J., & Ellison, A. M. (2013). *A primer of ecological statistics* (Second edition). Sinauer Associates, Inc., Publishers.
- Hastie T, Tibshirani R, Friedman J. 2016. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2 ed. Springer. [<https://web.stanford.edu/~hastie/ElemStatLearn/>].
- Hankin, D., Mohr, M. S., & Newman, K. B. (2019). *Sampling Theory: For the Ecological and Natural Resource Sciences* (1<sup>o</sup> ed). Oxford University Press.
- Healy K. 2019. *Data Visualization: a practical introduction*. Princeton University Press.
- Hector, A. (2021). *The New Statistics with R: An Introduction for Biologists* (2<sup>o</sup> ed). Oxford University Press.
- James G, Witten D, Hastie T, Tibshirani R. 2013. *An Introduction to Statistical Learning: with Applications in R*. 2.ed. Springer. [<http://faculty.marshall.usc.edu/gareth-james/ISL/>]
- Kabacoff RI. 2015. *R in Action: Data analysis and graphics with R*. 2.ed. Manning.
- Lander JP. 2017. *R for Everyone: Advanced Analytics and Graphics*. Addison-Wesley Professional.
- Manly, B. F. J., Navarro Alberto, J. A., & Brown, J. (2015). *Introduction to ecological sampling*.
- Matloff N. 2011. *The Art of R Programming: A Tour of Statistical Software Design*. No Starch Press.
- Oliveira PF, Guerra S, McDonnell, R. 2018. *Ciência de dados com R – Introdução*. IBPAD. [<https://cdr.ibpad.com.br/>]
- R Core Team. 2020. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. [<https://www.r-project.org/>]
- Scheiner, S. M., & Gurevitch, J. (2001). *Design and analysis of ecological experiments*. Oxford University Press.
- Teetor P. 2011. *R Cookbook*. O'Reilly Media. [<http://www.cookbook-r.com/>]
- Wickham H, Grolemund, G. 2017. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. O'Reilly Media. [<https://r4ds.had.co.nz/>]
- Wickham H. 2019. *Advanced R*. 2 ed. Chapman and Hall/CRC. [<https://adv-r.hadley.nz/>]
- Wickham H. 2020. *ggplot2: Elegant Graphics for Data Analysis*. 3 ed. Springer. [<https://ggplot2-book.org/>]
- Wilk CO. 2019. *Fundamentals of Data Visualization: A Primer on Making Informative and Compelling Figures*. O'Reilly Media. [<https://serialmentor.com/dataviz/>]

Zuur, A. F., Ieno, E. N., & Smith, G. M. (2007). *Analysing ecological data*. Springer.

Zuur, A. F., Ieno, E. N., Walker, N., Saveliev, A. A., & Smith, G. M. (2009). *Mixed effects models and extensions in ecology with R*. Springer New York.