



Relatório de Dados da Disciplina

Sigla: LCF5900 - 2 Tipo: POS
Nome: Processamento Reproduzível e Aberto de Dados Científicos
Área: Recursos Florestais (11150)

Datas de aprovação:

CCP: 08/09/2021 CPG: 05/10/2021 CoPGr:

Data de ativação: 05/10/2021 Data de desativação:

Carga horária:

Total: 120 h Teórica: 15 h Prática: 5 h Estudo: 10 h

Créditos: 8 Duração: 4 Semanas

Responsáveis: 85364 - Luiz Carlos Estraviz Rodriguez - 05/10/2021 até data atual

Objetivos:

O curso propõe um conjunto de procedimentos, e apresenta ferramentas, que permitem aos pesquisadores tratar os dados de forma aberta e reproduzível. O objetivo é desenvolver trabalhos científicos que permitam a qualquer pessoa compreender e replicar as etapas de uma análise, replicando-as novamente. A produção de resultados científicos reproduzíveis e abertos, e o uso de fluxos abertos de trabalho, permitem o compartilhamento de análises e a colaboração com outras pessoas, bem como a publicação aberta de dados e fluxos de trabalho para uma melhor disseminação do conhecimento científico.

Justificativa:

A ciência aberta envolve a disponibilização de métodos, dados e resultados científicos para todos. Em 2009, Dan Gezelter (<http://openscience.org/what-exactly-is-open-science/>) citou os seguintes atributos para a ciência aberta: transparência na coleta de dados; clareza na descrição dos métodos de processamento, análise e derivação de resultados; acesso público aos dados e métodos de processamento; transparência na comunicação dos resultados; e a colaboração facilitada por mecanismos web. Introduzimos neste curso os meios para que adesão a esses princípios e atributos. Este curso se inspira na forma como o Earth Lab da Universidade de Colorado, Boulder - EUA, forma novos pesquisadores.

Conteúdo:

1. Princípios do processamento reproduzível e aberto de dados; 2. Fluxos de trabalho científico reproduzível e aberto; 3. Organização de arquivos de dados; 4. Uso de Jupyter Notebooks; 5. Adoção do Git/GitHub para controle de versão; 6. Introdução ao Python; 6.1 Variáveis, listas, pacotes, instruções condicionais, loops e funções; 7. Introdução ao R; 7.1. Variáveis, vetores, pacotes, instruções condicionais, loops e funções; 8. R tidyverse para manipulação de dados; 9. R ggplot2 para geração de gráficos; 10. R markdown para relatórios de análise de dados; 11. R shiny para disponibilização de resultados na web.

Bibliografia:

Fernanda Peres (2021) Tutoriais sobre R. Disponível em: https://www.youtube.com/channel/UCx1vXgGDxc7BH7b11KC_o7Q. Acesso em 30 de jun. de 2021.
Finding Your Way To R. Disponível em: <https://education.rstudio.com/learn/>. Acesso em 30 de jun. de 2021.
How do you create a statistical model using tidymodels Disponível em: <https://www.tidymodels.org/start/>. Acesso em 30 de jun. de 2021.



Relatório de Dados da Disciplina

Leah Wasser & Jenny Palomino (2021) Introduction to Earth Data Science. Earth Lab at Colorado University, Boulder - USA. Disponível em: <https://www.earthdatascience.org/courses/intro-to-earth-data-science/>. Acesso em 30 de jun. de 2021.

Learn the tidyverse Disponível em: <https://www.tidyverse.org/learn/>. Acesso em 30 de jun. de 2021.

Lista de livros criados com o pacote bookdown Disponível em: <https://bookdown.org/>. Acesso em 30 de jun. de 2021.

Manual do pacote rmarkdown. Disponível em: <https://bookdown.org/yihui/rmarkdown-cookbook/>. Acesso em 30 de jun. de 2021.

Manual do pacote bookdown Disponível em: <https://bookdown.org/yihui/bookdown/>. Acesso em 30 de jun. de 2021.

Marcus Nunes (2021) Workshop R. Disponível em: <https://github.com/mnunes/workshopR>. Acesso em 30 de jun. de 2021.

Principais marcadores markdown guide. Disponível em: <https://www.markdownguide.org/cheat-sheet/>. Acesso em 30 de jun. de 2021.

Statistical Inference via Data Science: A Modern Dive into R and the Tidyverse. Disponível em: <https://www.tidymodels.org/books/moderndive/>. Acesso em 30 de jun. de 2021.

Yihui Xie, Christophe Dervieux & Emily Riederer (2021) R Markdown Cookbook. Disponível em: <https://bookdown.org/yihui/rmarkdown-cookbook/>. Acesso em 30 de jun. de 2021.

Forma de avaliação:

Média aritmética de dois trabalhos submetidos para avaliação coletiva quanto à reprodutibilidade e qualidade do relatório disponibilizado em nuvem

Observação:

A disciplina é apresentada de forma concentrada e pode contar com a participação de professor ou pesquisador internacional especialmente convidado como conferencista

Tipo de oferecimento da disciplina: Presencial
