

Projetos com R, Git e Github

Willber Nascimento

19/05/2022

Treinamento Brasileiro de Ciência Aberta

Este programa é para um curso de maior extensão, contudo ele pode guiar os participantes em seus estudos. Especificamente no Treinamento Brasileiro de Ciência Aberta teremos uma apresentação mais focada em apresentar as vantagens e os passos básicos do uso dessas ferramentas.

Ementa

Gerenciamento de projetos de análise de dados com R, Git e integrações com ferramentas de gerenciamento remoto como GitLab, GitHub e outros.

O que você deveria conhecer

Para tirar o máximo proveito do material, assumimos uma certa familiaridade básica com R. Por básico entendemos: instalar o software, criar, importar e exportar arquivos. Você também deve saber como instalar um pacote no R e o carregar na sessão. Portanto se você sabe o que é `install.packages`, `library`, `read.csv` você conseguirá aproveitar completamente curso. **Mas se você não sabe, a apresentação pode te mostrar como iniciar sua jornada com as ferramentas. Insisto que você participe.**

Além disso, assumo que você sabe como instalar programas no seu computador. Os conceitos e ferramentas tratadas no curso são, em sua maioria, *open source* e estão disponíveis em todas as plataformas (Linux, Mac OS e Windows).

Objetivos

Esperamos que ao final do curso os participantes (1) entendam a importância do gerenciamento de projetos na prática da análise de dados, transparência, replicabilidade e trabalho colaborativo, bem como (2) tenham conhecimentos mínimos para implementar na prática ferramentas de gerenciamento de projetos de forma moderna com R e Git.

Livro texto

Neste curso usaremos múltiplas fontes, mas se eu pudesse recomendar uma referência geral seria Gandrud (2013). Com ele você aprende as principais ferramentas disponíveis para tornar seu trabalho reprodutível utilizando o R e o RStudio, mas ele discute conceitos que também podem ser utilizados em qualquer plataforma.

Programação e Plano de Aula

1. Fluxo de trabalho baseado em projetos

Vamos discutir e apresentar argumentos para implementação de um fluxo de trabalho baseado em projetos. Para esse tópico recomendamos que você acesse:

- Leia o capítulo 8 de Wickham and Grolemond (2016)
- Veja o post da Bryan (2017) no blog do Tidyverse

2. Instalação e configuração das ferramentas (R, Git, Github/lab)

Demonstração de como instalar as principais ferramentas de gerenciamento de projetos que serão utilizadas no curso.

- Exemplificaremos esse tópico usando o Linux. Você pode optar por seguir algum material na internet usando seu sistema operacional.
- Acessem esses videos para instalarem os programas
 - Como instalar: <https://youtu.be/wE5fWsPHcps> e <https://youtu.be/H9MXEuYIMDc>
 - Como configurar: <https://youtu.be/zOG0LtOyjNA>

3. Introdução ao controle de versão: noções de Git e Github

Apresentação dos principais conceitos abordados no controle de versão de arquivos e demonstração da utilidade do controle de versão usando Git e Github.

- Vamos apresentar os principais conceitos dos capítulos 1 e 2 do Chacon and Straub (2014)
- Uma fonte importante de consulta são as documentações do Github e Gitlab.

4. Integração do R com Git e Github

Exemplifica como integrar controle de versão com R e RStudio, bem como suas principais funções.

- Consulte o Capítulo 5 de Gandrud (2013) como referência básica, mas fique a vontade para usar qualquer material disponível na internet.

Material e Leituras

Softwares utilizados

- R e Rstudio
- Git e Github
- Um navegador de internet moderno

Leituras recomendadas

Bryan, Jenny. 2017. “Project-Oriented Workflow.” <https://www.tidyverse.org/blog/2017/12/workflow-vs-script/>.

Chacon, Scott, and Ben Straub. 2014. *Pro Git*. Springer Nature.

Frank, Michael C, and Rebecca Saxe. 2012. “Teaching Replication.” *Perspectives on Psychological Science* 7 (6): 600–604.

Gandrud, Christopher. 2013. *Reproducible Research with r and r Studio*. CRC Press.

- Janz, Nicole. 2016. "Bringing the Gold Standard into the Classroom: Replication in University Teaching." *International Studies Perspectives* 17 (4): 392–407.
- King, Gary. 1995. "Replication, Replication." *PS: Political Science and Politics* 28 (3): 444–52.
- Wickham, Hadley, and Garrett Grolemund. 2016. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. " O'Reilly Media, Inc."