

Introdução ao R

Iara Passos

iaracpassos@gmail.com

Este curso pretende apresentar uma introdução a uma das ferramentas de análise de dados mais versáteis e completas em uso atualmente: a linguagem e ambiente de programação R. Iremos, ao longo de três encontros, aprender como utilizar suas funcionalidades básicas de forma rápida e reprodutível com a intenção de utilizá-las e aplicá-las na pesquisa científica. Para tanto, após a compreensão das ferramentas básicas da linguagem, iniciaremos uma introdução a análise de dados, construção de gráficos e a exportação deste material.

Cabe ressaltar que a intenção deste curso não é de esgotar as possibilidades do ambiente R, mas sim introduzir a linguagem e capacitar os/as alunos/as a adaptar o conhecimento adquirido aos seus problemas e finalidades de pesquisa. Ao final do curso, espera-se que os/as participantes estejam familiarizados com a sintaxe do R e estejam capacitados para entender qualquer código escrito em R e saber como procurar novas funcionalidades da linguagem na documentação disponibilizada em *websites*, *blogs*, livros e etc.

Local de realização do curso: Mconf

Dias: 10/10 17/10 24/10 (sábados)

Horário: 8h30-12h30

Carga Horária Total: 12h

Repositório: Gitlab

Pré-requisitos

Não é necessário nenhum conhecimento prévio de programação para a realização desse curso. Conhecimentos básicos de estatística descritiva são desejáveis.

É necessário o cadastro prévio no site do RStudio Cloud.

Funcionamento do curso

As aulas serão práticas. O desenvolvimento das atividades será realizado com compartilhamento de tela da ministrante de forma que os/as participantes possam acompanhar a execução e as reproduzir em seus próprios equipamentos. Em algumas situações poderão ser utilizados *slides*. Há também a possibilidade de recomendação de leitura de material prévio para as aulas, além do envio de lista de exercícios para fixação dos conceitos trabalhados em

aula. O estudo deste material não é obrigatório, mas é altamente recomendado para melhor desenvolvimento do curso.

Caso a/o cursista precise de ajuda com o código, poderá ser solicitado que passe o código via Pastebin ou compartilhe a sua tela.

Todo o material do curso (*scripts*, *slides*, programa etc) será disponibilizado. Para melhor funcionamento das aulas, os *scripts* e exercícios da aula correspondente serão disponibilizados após a finalização da mesma.

Caso seja necessário, auxílio e elucidação de dúvidas entre uma aula e outra podem ser feitos por e-mail.

Material de apoio

Além do material do curso, que será disponibilizado, são recomendados os seguintes materiais:

Livros

R for Data Science de Garrett Golemund & Hadley Wickham

Hands-On Programming with R de Garrett Golemund

ggplot2: Elegant Graphics for Data Analysis de Hadley Wickham

Dúvidas em geral

Stack Overflow

RStudio Cheatsheets

RStudio Primers

R Bloggers

Bookdown

Listas de exercícios

A realização dos exercícios para esse curso não são obrigatórios, mas é fundamental para melhor fixação das atividades, sobretudo pelo fato das dúvidas que tendem a surgir justamente durante a realização dos exercícios. Os exercícios serão disponibilizados após a finalização de cada aula em arquivo formato .R. Caso tenham dúvidas, o/a participante deve encaminhar a lista dos exercícios no mesmo arquivo (script) por email e escrever no próprio script em

forma de comentário no exercício correspondente. Alguns exercícios podem ser discutidos na aula subsequente, caso haja demanda.

Regras básicas

1. Ninguém fica para trás

Para que isso seja possível, é muito importante que perguntas sejam feitas durante a aula. Também conto com a colaboração dos/as participantes que tenham mais facilidade ou que já tenham compreendido o tópico em questão para que ajudem os/as colegas que estejam com dificuldade em algum ponto.

2. Atenção e concentração

O curso segue uma sequência de conceitos e ferramentas que dependem do tópico anterior. Por isso, é necessário muita atenção ao longo de todas as aulas, principalmente durante a explicação.

3. Pontualidade

Para melhor funcionamento das atividades as aulas irão começar rigorosamente às 8h30.

4. Ética

A reprodução integral de qualquer *script* e código de terceiros sem a devida referência é considerado plágio, em qualquer local de reprodução. Ainda que não estejam sendo avaliados para nota e afins é importante salientar essa questão ética.

Plano das aulas

Aula 1 Introdução

Data: **24/10**

1. Apresentação
2. Conhecendo o ambiente
3. Pacotes e ajuda
4. Primeiros passos
5. Objetos e classes de objetos
6. Vetores e matrizes
7. Funções e factors

Aula 2 Manipulando dados

Data: 17/10

1. Dataframes
2. Listas
3. Importando bancos de dados
4. Tratamento inicial dos dados
5. Estatísticas descritivas
6. Exportando bancos de dados

Aula 3 Visualização dos dados

Data: 24/10

1. *Tidyverse*
2. Gráficos básicos no R
3. O pacote *ggplot2*
4. Exportando gráficos
5. Markdown e RMarkdown