

Bibliografia e material para o curso de apoio de R

Professores: Felipe Cornélio e Marcelo Tonon

Atualizado em 04/06/2020

Objetivo do Curso

O objetivo deste curso é capacitar os alunos da disciplina “Tópicos especiais em técnicas insumo-produto: teoria, fundamentos e aplicações”, do PPGE/IE, ministrada pelos professores Esther Dweck e Kaio Vital da Costa, nas habilidades básicas de programação por meio do software R. As atividades de monitoria terão um caráter prático de aplicação de tópicos discutidos em sala, de modo a que os estudantes desenvolvam habilidades necessárias ao desenvolvimento de seus trabalhos acadêmicos. As atividades serão desenvolvidas especificamente para manipulação de matrizes de insumo-produto. O curso será dividido entre as “Aulas Fundamentais”, as “Aulas Práticas” e as “Aulas Suplementares”.

Habilidades fundamentais

O curso, apesar de ser voltado para matrizes, se propõe a introduzir um mínimo de capacidades básicas para os alunos, listadas abaixo:

1. Entender as diferentes estruturas de dados
2. Saber importar e exportar dados
3. Saber automatizar tarefas repetitivas
4. Saber usar condicionantes como o if
5. Saber manipular e extrair Subconjuntos de vetores (por linhas, colunas, valores e caracteres)
6. Realizar operações matemáticas
7. Criar funções
8. Realizar gráficos
9. Ser relativamente independente para resolver erros e aprender novos conceitos.

Essas habilidades serão o foco principal das aulas fundamentais. É de certo modo óbvio que 1 hora semanal de aula assistida, não é o suficiente para solidificar essas habilidades. Assim sendo, se faz necessário um empenho individual em separado no aprendizado, de modo que é encorajado buscar pratica-las. Dessa forma, o esforço em superar, e entender, as dificuldades e os erros, que inevitavelmente aparecerão ao longo do processo são extremamente necessários para a consolidação destas habilidades.

Bibliografia

Em boa parte das aulas, o curso terá um material próprio focado para o aplicação do R na Matriz Insumo Produto (MIP). Boa parte deste próprio material é baseado e inspirado em outros materiais/cursos. Materiais esses que tem uma visão bastante ampla sobre o assunto o

que não é tão interessante assim dado o escopo do curso. Outros materiais vão explicar a mesma coisa só que de forma diferente. De toda forma, pressupomos que todos os materiais direcionados (MD) serão lidos.

Destaque: Hadley Wickham

Muito do material que cuja produção foi direcionada especialmente para o curso, vem de livros feitos e disponibilizados de forma online e gratuita pelo Hadley Wickham(HW), que é muito conhecido pela criação dos seus pacotes que são coletivamente chamados de *tidyverse*. Dentro da produção, três livros foram basilares na hora da concepção do material direcionado a MIP:

1. **Hands-on Programming with R (HOPR):** Publicado em 2014, é um livro bastante introdutorio, voltado para ser muito amigável quem não entende nada de programação. O aprendizado é cadenciado e o autor aqui faz tudo a partir de exemplos. Vale a pena dar uma lida especialmente para quem nunca usou o R, mas também para os usuários mais antigos vale uma foliar ao menos o sumário, pois pode haver tópicos pouco familiarizados pelos usuários, como *Debugging* por exemplo.
2. **R for Data Science (R4DS):** Publicado em 2017, é já um livro que considera alguma, mas não muita, intimidade com o R e o Rstudio. O foco principal do livro é ensinar o usuário a mexer com os dados usando os pacotes do *tidyverse*. Para quem já tem alguma intimidade com programação e/ou o R, esse livro pode ser o mais indicado.
3. **Advanced R (ADV-R):** A primeira edição deste livro saiu em 2014, e a segunda em 2019. O livro é já indicado para quem tem alguma vivência no R, especialmente do capítulo 10 (2 ed.) em diante. Até então ele é mais complexo, porém tem a vantagem de dar profundidade e amplitude a certos conceitos. Mesmo para quem não é experiente, ele pode ser de ajuda quando se precisa ver o quadro mais geral para entender um aspecto.

Capítulos e seções destes três livros serão colocados como referência para cada aula, para quem quiser ir além do material direcionado a MIP. A sugestão é de que se leia primeiramente o **R4DS** e então se avalie se é necessário dar um passo atrás e conferir o **HOPR** ou buscar ampliar o conhecimento e se aventurar pelo **ADV-R**. Sempre que houver um *, é sinal de que a leitura deste capítulo ou seção, indicados dentro do parenteses, é bastante proveitosa e bastante indicada para o tema.

De todo modo, se para quem quiser ter um conhecimento relativamente profundo e amplo, o que está fora do escopo deste curso, vale a leitura completa dos livros do Hadley Wickham, que são além de tudo um material gratuito que está sempre sendo atualizado.

Cronograma e Bibliografia das aulas

Links para as fontes externas estão no final. Todas em **negrito** são consideradas fundamentais para o andamento da monitoria.

03/06/2020 - Aula Fundamental 0 - Instalação e Introdução ao R e o RStudio.

Conteúdo: Apresentação dos propósitos do curso, auxílio em eventuais dificuldades com a instalação dos softwares e a demonstração de um pouco de como mexer no Rstudio.

MD(aula0&1), ADV-R (1*) e R4DS(1.6).

10/06/2020 - Aula Fundamental 1 - Vetores e Operações básicas

Como se escreve no R. – **MD(aula0&1)**, HOPR(2).

Introdução aos tipos mais comuns dos Vetores e as estruturas dos Dados – **MD(aula0&1)**, HOPR(5.1-8), R4DS(20), ADV-R(3), FB(DTS).

Operações básicas do R - **MD(aula0&1)**, **Algebra Linear**.

17/06/2020 - Aula Fundamental 2 - Leitura e exportação de dados (Parte 1)

Funções Específicas para a leitura dos dados de diferentes tipos bases (.Rdata, .csv e excel) e os desafios que podem emergir em cada uma.

MD(aula2), HOPR(Apêndice B*, 5.9-11), FB(ID), SW(11).

24/06/2020 - Aula Fundamental 3 –Condicionantes e Limpeza de dados

if e *ifelse* - **MD(aula3)**, HOPR(9), ADV-R(if=5.1-2).

Introdução ao *subsetting* pelo *base-R* –**MD(aula3)**, HOPR(5-6), ADV-R(4).

Introdução ao *subsetting* pelo *tidyverse*: R4DS(5, 12-13).

01/07/2020 - Aula Fundamental 4 -Automatizando tarefas

for e as variações da função *map*

MD(aula4), HOPR(11-12), R4DS(21.1-3), ADV-R(5.3, 9), SW(15).

08/07/2020 - Aula Fundamental 5 - Leitura e exportação de dados - parte 2

Função *list.files()* e pacote *stringr* – **MD(aula5)**, R4DS(14), SW(13).

Importação automática da base de dados a partir de múltiplos arquivos – MD(aula5) e R4DS(11).

15/07/2020 - Aula Fundamental 6 - Introdução ao ggplot2.

Produção de gráficos a partir do *ggplot2*. Discussão: quando fazer e quando não fazer?

MD(aula6), R4DS(3, 28), Livro do *ggplot2* feito pelo HW.

22/07/2020 - Aula Fundamental 7 - Funções e Aula com exemplo

Discussão sobre funções - **MD(aula7)**, HOPR(2), R4DS(19), ADV-R(6-8), SW(15).

Aula com um exemplo prático.

Possíveis Aulas Suplementares:

Aula Suplementar – Programação Funcional pelo pacote purrr

O que é Programação Funcional e como utilizar as funções da família map do pacote **purrr**

MD(Sup), R4DS(21), ADV-R(9-11)

Aula Suplementar – Aprofundando em Listas

Aula Suplementar – tabulizer, word e LaTeX

Pacotes para importação de tabelas para o R direto do PDF e para exportar para Word e LaTeX.

Aula Suplementar – Tidyverse

Introdução aos vários pacotes do Tidyverse e quais são em especial uteis para nós. Pela extensão do tidyverse e o pouco tempo da aula, não entraremos em muitos detalhes em cada tema.

Aula Suplementar – Boas Práticas

Apresentar o conceito de Boas Práticas de programação e um guia para seguir.

Aula Suplementar – Git e Github

Como utilizar o Git e o Github nos seus projetos do RStudio. - LivroGIT&R

Referências:

Obrigatórias:

MD: “Material Direcionado” feito pelos monitores especialmente para o curso e focados no uso em Matriz Insumo Produto. Serão entregues conforme o andamento do curso.

Intro_AL: “Introdução à Álgebra linear com R” feito pelo Kaio Vital.

Livros do Hadley Wickham

HOPR: <https://rstudio-education.github.io/hopr/>

R4DS: <https://r4ds.had.co.nz/>

ADV-R: <https://adv-r.hadley.nz/>

Livro do ggplot2 feito pelo Hadley Wickham: <https://ggplot2-book.org/>

Outras fontes:

SW: <https://swcarpentry.github.io/r-novice-inflammation/>

É digno de nota que o conteúdo do SW ficou um pouco datado em alguns poucos pontos, especialmente por conta da atualização do R na versão 4.0.0. A parte 11, indicada para a aula 2, sofre desse ponto. Porém de modo geral esse curso é uma boa alternativa gratuita para quem quiser aprender.

FB(ID): <https://felipegalvao.com.br/blog/2015/11/16/r-basico-importacao-de-dados-para-o-r/>

FB(DTS): <http://felipegalvao.com.br/blog/2015/11/04/r-basico-introducao-tipos-e-estruturas-de-dados-2/>

LivroGIT&R: <https://happygitwithr.com/>

Purrr: https://www.brodrigues.co/blog/2017-03-24-lesser_known_purrr/

Extras

- Página do Rstudio onde tem vários pdfs com “colas” para várias coisas dentro do R:

<https://rstudio.com/resources/cheatsheets/>

Para tirar dúvidas e aprender mais, veja *1.6 Getting help and learning more* no R4DS.