

# Introdução ao R

**Fernando Meireles**

Docente

## Ementa

Este curso oferece uma introdução a uma das ferramentas mais potentes para análise de dados: o ambiente de programação **R**. Entre outros, o curso cobrirá desde a coleta e manipulação de dados até a produção de estatísticas descritivas e visualizações. As aulas serão acompanhadas de atividades práticas que envolverão extrair e limpar dados de diversos meios, sintetizar informações e exportar resultados de forma automatizada para *Word*,  $\text{\LaTeX}$ , *PDF* e *HTML*.

O curso está organizado em quatro aulas. Na primeira, trabalharemos com o básico da sintaxe em **R** necessário para criar e manipular vetores e bases de dados (*data.frames* e *tibbles*). Já nas aulas 2 e 3, veremos como utilizar pacotes do *tidyverse* para realizar tarefas de limpeza, manipulação, visualização e síntese de dados. Por fim, na aula final aplicaremos estes conhecimentos – além de outros que usaremos para documentar, exportar e tornar replicável os resultados de um projeto – para realizar tarefas práticas, comuns em análises de dados.

Certamente não esgotaremos as possibilidades do mundo da análise de dados com o **R** (na verdade, seria irrealista prometer isto). Ao fim do curso, entretanto, é esperado que os(as) alunos(as) estejam prontos para incorporar o **R** as suas rotinas de trabalho.



## Programa

### Aula 1 Básico

1.1 Introdução ao R; 1.2 Objetos, vetores e classes; 1.3 `data.frame`; 1.4 Indexação e manipulação de `data.frame`.

### Aula 2 Manipulação de dados

1.1 Funções e pacotes; 1.2 Introdução ao *tidyverse*; 1.3 Leitura de dados; 1.4 Manipulação de `data.frame` com `dplyr`.

### Aula 3 Visualização

1.1 Manipulação de dados II; 1.2 Síntese de informações; 1.3 Criação de gráficos com *ggplot2*; 1.4 Estatística descritiva e modelos de regressão.

### Aula 4 Exportação

1.1 Introdução ao R Markdown; 1.2 Exportação de gráficos e tabelas; 1.3 Criação de documentos.

## Pré-requisitos

Para este curso, será necessário trazer seu próprio *notebook* para acompanhar as aulas. Além disso, é importante já ter o R e outros *softwares* que precisaremos.

Para instalar o R, basta ir ao [site do CRAN \(Comprehensive R Archive Network\)](#), que é a rede de fundadores e administradores do core da linguagem R, e baixar o setup indicado para o seu sistema operacional.



Feito isto, já é possível usar o R – mas só via console, o que não é tão fácil/útil. É por isso que usaremos uma IDE (i.e. Ambiente de Desenvolvimento Integrado) neste curso: especificamente, usaremos o RStudio. Para baixá-lo, basta [entrar neste site](#) e escolher a opção mais adequada para o seu sistema operacional. Para um tutorial detalhado, [este guia pode ser útil](#).

## Referências

- Wickham, H.; Grolemund, G (2017). [R for Data Science](#). O'Reilly.
- Meireles, F.; Silva, D. (SD). [Usando R: Um Guia para Cientistas Políticos](#). Manuscrito.



## Docente

### Fernando Meireles

Professor Adjunto de Ciência Política do Instituto de Estudos Sociais e Políticos da Universidade do Estado do Rio de Janeiro (IESP-UERJ) e Doutor em Ciência Política pela Universidade Federal de Minas Gerais (UFMG). Trabalha nas áreas de Política Comparada, Instituições Políticas e métodos quantitativos aplicados às Ciências Sociais, com ênfase em aprendizado de máquina e inferência causal.

