[Análise de dados com R  
Da introdução à autonomia intermediária](https://www.sympla.com.br/evento-online/analise-de-dados-com-r-da-introducao-a-autonomia-intermediaria/2657807)

O curso trata de análise de dados com a linguagem de programação R, nos ambientes do RStudio/Posit e Google Colab. O conteúdo é voltado para operações fundamentais visando o domínio e a autonomia intermediária no uso do software para análise de dados, estruturado seguindo os  eixos do livro [***R for Data Science***](https://r4ds.hadley.nz/), com foco em conteúdo prático, com tutoriais e exercícios. Envolve desde operações básicas de importação e manuseio de dados até a análise, visualização e comunicação.

**Sobre a ferramenta**

A linguagem **R** é uma das principais ferramentas no universo da **Ciência de Dados**, amplamente adotada em instituições de pesquisa, órgãos públicos e empresas ao redor do mundo. Seu código aberto e uso gratuito, disponibilizado através da [*The Comprehensive R Archive Network (r-project.org)*](https://cran.r-project.org/#:~:text=R%20is%20%E2%80%98GNU%20S%E2%80%99,%20a%20freely%20available), garantem flexibilidade e acessibilidade. Além disso, é possível de ser integrado aos ambientes de desenvolvimento *[RStudio Desktop - Posit](https://posit.co/download/rstudio-desktop/" \t "_blank)* e **Google Colab**, facilitando o desenvolvimento de projetos colaborativos e replicáveis. Com uma vasta comunidade de usuários e desenvolvedores, o R conta com uma biblioteca robusta de pacotes para **análise, visualização de dados e aprendizado de máquina**, sendo uma escolha cada vez mais popular por sua eficiência no tratamento de grandes volumes de dados e pela capacidade de gerar visualizações ricas e informativas.

**Dinâmica**

O curso está organizado em cinco sessões online síncronas, com duração de 1 hora e 30 minutos. As sessões ocorrerão às quintas-feiras das 20h30 às 22h00, e serão gravadas para aqueles que não puderem participar, disponibilizadas por até uma semana.

**Público-alvo**

O curso é direcionado a pessoas que trabalham com pesquisa acadêmica e de mercado, bem como analistas e gestores que utilizam dados primários e/ou secundários.  
  
**Materiais**

Todos os materiais serão disponibilizados em pasta compartilhada pelo Google Drive e em um repositório do [GitHub](https://github.com/victorgalcantara/R-Introduction).

**Instrutor:** Victor Gabriel Alcantara é licenciado em Ciências Sociais e mestre em Sociologia (PPGSA/UFRJ), com MBA em Data Science (USP/ESALQ). Atualmente está cursando o doutorando em Sociologia (PPGS/USP). Tem experiência com pesquisa acadêmica e de mercado, consultorias estatísticas para pesquisadores e organizações do terceiro setor. Saiba mais: <https://linktr.ee/victorgalcantara>.

**Inscrições até 10/out ([clique para se inscrever](https://www.sympla.com.br/evento-online/analise-de-dados-com-r-da-introducao-a-autonomia-intermediaria/2657807))**

**Período do curso:** 31 de outubro à 28 de novembro.

|  |  |
| --- | --- |
| Conteúdo Programático | |
| **1ª Aula**  **(1 h 30 min) 31/10** | **Explore: basic knowledge and workflow**  **Noções gerais: é de comer?**  • Apresentação R e RStudio;  • Fluxo de trabalho no RStudio: janelas, linguagens suportadas e convenções para o desenvolvimento de códigos;  • Práticas comunitárias: help, Stack Overflow e GitHub.  **Tutorial 01:** instalação e setup. |
| **2ª Aula**  **(1 h 30 min) 07/11** | **Explore: logic operations and basic R**  **Mão na massa**  • Operações básicas, comandos e lógica  • Classes de objetos: vetores, matrizes, listas e data.frames  • Funções e pacotes: o universo do tidyverse  **Tutorial e exercícios 02:** operações básicas e aplicação com dados. |
| **3º Aula**  **(1 h 30 min)**  **14/11** | **Import and tidy data**  **Dados estruturados**  • Import: importação de dados  • Tidy and Transform: manuseio dos dados (filter, select, rename, mutate, group by etc.)  **Tutorial e exercícios 03:** manuseio de dados. |
| **4ª Aula**  **(1 h 30 min)**  **21/11** | **Analyse**  **Estatísticas descritivas**  - Tabela de frequência  - Medidas de tendência central (média e mediana)  - Medidas de dispersão (amplitude, variância e desvio-padrão)  **Tutorial e exercícios 03:** análise descritiva de dados. |
| **5ª Aula**  **(1 h 30 min)**  **28/11** | **Visualize**  **Gráficos**  Visualização de dados com GGplot: A Grammar of Graphics  - Barras e Pizza  - Barras, boxplot, histogramas e densidade  - Cartografias  **Exercício final:** relatório com análise de um dado que tenha interesse. |
| Referências | |
| Wickham, H. & Grolemund, G. ***R for Data Science***(O’Reilly, Canadá, 2017).  Ihaka, R. & Gentleman, R. R: **A Language for Data Analysis and Graphics**. *Journal of Computational and Graphical Statistics* **5,** 299. issn: 10618600. https://www.jstor.org/stable/1390807?origin=crossref (2023) (set. de 1996).  Ihaka, R. **R : Past and Future History**. A Draft of a Paper for Interface ’98 (1998).  Angrist, J. D. & Pischke, J.-S. ***Mastering metrics****: the path from cause to effect* 282 pp. isbn: 978-0-691-15283-7 978-0-691-15284-4 (Princeton University Press, Princeton ; Oxford, 2015).  Imai, K. ***Quantitative social science****: an introduction* OCLC: ocn958799734. 408 pp. isbn: 978-0-691-17546-1 978-0-691-16703-9 (Princeton University Press, Princeton, 2017).  Wickham, H. *GGplot2:* ***Elegant Graphics for Data Analysis***(Springer International Publishing, New York, 2016) | |