

PROPOSTA DE OFICINA DIDÁTICA NO RSTUDIO: UMA INTRODUÇÃO AO PACOTE GGLOT2

ISADORA MOREIRA DA LUZ REAL¹; ANA LUIZA BARBOZA MERLIN²;
FERNANDO NEUGEBAUER REHBEIN DA CUNHA PENEDO³; GISELDA MARIA
PEREIRA⁴; POLLYANE VIEIRA DA SILVA⁵

ANA RITA DE ASSUMPÇÃO MAZZINI⁶:

¹Universidade Federal do Rio Grande – isadora.real18@hotmail.com

²Universidade Federal de Pelotas – merlinanaluiza@gmail.com

³Universidade Federal de Pelotas – fnrda@inf.ufpel.edu.br

⁴Universidade Federal de Pelotas – gmpereira08@gmail.com

⁵Universidade Federal de Pelotas – pollyane.silva@ufpel.edu.br

⁶Universidade Federal de Pelotas – anarita.mazzini@gmail.com

1. INTRODUÇÃO

As ferramentas computacionais para a estatística e ciência de dados vem se tornando mais robustas ao passar dos anos (WICKHAM; GROLEMUND, 2017). A linguagem R surgiu no início dos anos 90, criado por Ross Ihaka e Robert Gentleman da Universidade de Auckland, na Nova Zelândia (CRAWLEY, 2012), com foco em análises estatísticas.

Por ser uma linguagem livre e de código aberto, angariou diversos colaboradores e vários pacotes para visualização gráfica de dados surgiram (MORALES, 2025; R CORE TEAM, 2025). Um pacote que ganhou destaque foi o ggplot2, criado por Hadley Wickham em 2005 (WICKHAM, 2016), como uma implementação dos princípios descritos no livro *The Grammar of Graphics*, de Leland Wilkinson, publicado no mesmo ano (WILKINSON, 2011). Neste pacote a visualização gráfica dos dados pode ser dividida em componentes semânticos, como escalas e camadas.

Em virtude disso, o ggplot2 permite ampla customização gráfica, possibilitando a construção desde gráficos simples, até visualizações mais complexas, como mapas, gráficos animados e composições que combinam diferentes tipos de gráficos em uma única visualização (WICKHAM, 2016). Além disso, o ggplot2 é compatível com diversos outros pacotes do ecossistema R, como o dplyr (WICKHAM et al., 2025), para manipulação de data frames; o plotly (PLOTLY TECHNOLOGIES INC, 2015), para criação de gráficos interativos; e o forcats (WICKHAM, 2023), voltado para a ordenação e manipulação de variáveis categóricas, entre outros.

Dessa forma, o domínio e o uso adequado de ferramentas para a elaboração gráfica de bancos de dados constituem uma habilidade cada vez mais valorizada tanto no mercado de trabalho quanto no meio acadêmico (KNAFLIC, 2017; WICKHAM; GROLEMUND, 2017). A visualização de dados por meio de gráficos permite não apenas descrever de forma clara e objetiva os resultados de uma pesquisa, mas também apresentar cronologias de eventos e padrões complexos de maneira acessível ao público-alvo (KNAFLIC, 2017). Assim, o uso eficiente de representações gráficas utilizando o pacote ggplot2 contribui para a comunicação científica e para a tomada de decisões baseada em dados.

Diante disso, o objetivo deste trabalho é descrever uma proposta de oficina didática com a utilização do pacote ggplot2 para alunos de graduação da Universidade Federal de Pelotas.

2. ATIVIDADES REALIZADAS

O desenvolvimento da proposta da oficina “Introdução ao ggplot2” faz parte das atividades do Projeto de Ensino “Ferramentas Computacionais no Ensino de Estatística” da Universidade Federal de Pelotas (UFPEL).

A oficina foi elaborada e pensada para alunos de graduação que já cursaram disciplinas de Estatística, com a finalidade de introduzi-los à sintaxe do ggplot2 para que possam utilizar na suas práticas profissionais ou acadêmicas. O tempo previsto para a oficina é de três horas, sendo separado em três dias. Onde no primeiro dia será a introdução e a familiarização dos participantes à sintaxe ggplot2. Segundo dia serão apresentados diferentes tipos de gráficos e como alterar as características dos mesmos. O terceiro e último dia será dedicado a uma breve demonstração de aplicações mais avançadas do ggplot2, incluindo a criação de mapas, gráficos animados e gráficos interativos. Também serão disponibilizados exercícios com outros conjuntos de dados, destinados à prática fora da oficina. Para fins de avaliação, será disponibilizado um formulário do Google, no qual os participantes poderão registrar suas opiniões sobre a oficina ministrada. O objetivo é verificar se a atividade atendeu às expectativas e identificar possíveis melhorias para futuras edições.

A concepção do material desta oficina surgiu a partir das demandas identificadas nas fichas de avaliação das oficinas anteriores, que evidenciaram o interesse dos participantes em aprender a utilizar o pacote ggplot2. Diante disso, com base nas oficinas anteriores desenvolvidas nos projetos “Explorando R como Ferramenta auxiliar nas disciplinas de Estatística” e “Projeto Ferramentas Computacionais no Ensino de Estatística”, foram selecionados alguns tipos de gráficos, os quais foram adaptados para a sintaxe do pacote ggplot2, na linguagem R (Figura 1). A partir dessa seleção, foi elaborado um script em R utilizando três bases de dados: o conjunto iris, disponível nativamente no R (R CORE TEAM, 2025); dados históricos de pluviometria da cidade de Pelotas, obtidos por meio do portal WorldClim(FICK; HIJMANS, 2017); e o conjunto gapminder, fornecido pelo pacote gapminder (BRYAN, 2025).

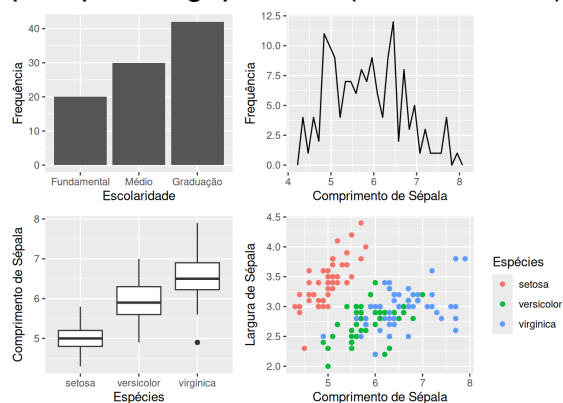


Figura 1- Exemplos de gráficos elaborados para a oficina “Introdução ao ggplot2”.
Fonte: Autor.

A partir do script e da estrutura da oficina, foi desenvolvido um site com o objetivo de disponibilizar de forma acessível o conteúdo instrucional sobre visualização de dados utilizando o ggplot2, servindo como material de apoio e consulta futura aos participantes. Para isso, empregou-se a ferramenta Quarto (ALLAIRE et al., 2025), que permite a integração de código, texto e visualizações interativas em documentos reproduzíveis.

A estrutura do site foi construída com arquivos no formato .qmd (Quarto Markdown), organizados em seções temáticas: introdução ao ggplot2, tipos de gráficos, personalização de elementos (cores, escalas, fontes, tamanhos e temas), combinação de múltiplos gráficos em uma única imagem, além de um módulo extra com conteúdos complementares (Figura 2). O conteúdo foi desenvolvido no ambiente RStudio (POSIT TEAM, 2025), com exemplos práticos e código comentado.



Figura 2 – Layout do site gerado via Quarto com o material didático para consulta da oficina “Introdução ao ggplot2”. Fonte: Autor.

Para a publicação, utilizou-se o sistema de controle de versões *Git*, aliado ao *GitHub*, com a hospedagem gratuita feita via *GitHub Pages*. A compilação dos arquivos e a geração do site estático foram realizadas com o comando *quarto publish*. Além do site, também foi gerada uma apresentação em PDF, por meio do próprio Quarto, para uso durante a oficina.

O site também disponibiliza para download a apresentação utilizada na oficina, o script completo em R, o arquivo shapefile da cidade de Pelotas e o raster contendo a média anual de precipitação no município, referente ao período de 1970 a 2000. Serão desenvolvidos exercícios complementares e disponibilizados no site, com o objetivo de permitir que os participantes pratiquem os comandos apresentados durante a oficina utilizando outros conjuntos de dados.

Todo o processo seguiu princípios de reprodutibilidade e ciência aberta, promovendo o acesso livre ao conteúdo por participantes e demais interessados em explorar recursos gráficos no R com o ggplot2.

3. CONSIDERAÇÕES FINAIS

O desenvolvimento deste material é importante para aplicação de futuras oficinas que abordem ggplot2, visto que é um conteúdo requisitado pelos participantes de oficinas anteriores dos Projetos de Ensino. Acredita-se que a

estrutura modular do conteúdo e a sua disponibilização online favoreça a autonomia dos participantes e o uso posterior como material de apoio.

A oficina tem previsão de realização no segundo semestre de 2025 e, espera-se que atenda as necessidades dos participantes, promovendo a integração dos conteúdos abordados, de modo que estes possam aplicá-los em seus próprios trabalhos e projetos.

4. REFERÊNCIAS BIBLIOGRÁFICAS

ALLAIRE, J; TEAGUE, C.; SCHEIDEGGER, C.; XIE, Y.; DERVIEUX, C.; WOODHULL, G. **Quarto** (Version 1.7), 2025.

BRYAN, J. **gapminder**: Data from Gapminder. R package version 1.0.1, 2025. Acessado em: 25 de julho de 2025. Disponível em: <https://CRAN.R-project.org/package=gapminder>

CRAWLEY, M. J. **The R Book**. John Wiley & Sons, 2012.

FICK, S. E.; HIJMANS, R. J. WorldClim 2: new 1km spatial resolution climate surfaces for global land areas. **International Journal of Climatology** v. 37 , n.12, p. 4302-4315, 2017.

KNAFLIC, C.N. **Storytelling with data**: A data visualization guide for business professionals. John Wiley & Sons, 2015.

MORALES, M. **sciplot**: Scientific Graphing Functions for Factorial Designs.R package version 1.2-0, 2025. Acessado em: 25 de julho de 2025. Disponível em: <https://github.com/mutualism/sciplot>.

PLOTLY TECHNOLOGIES INC. **Collaborative data science Publisher**: Plotly Technologies Inc. Montréal, QC, 2015. Acessado em: 25 de julho de 2025. Disponível em: <https://plot.ly>.

POSIT TEAM. RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA, 2025. Acessado em: 25 de julho de 2025. Disponível em: <http://www.posit.co/>.

R CORE TEAM. **R**: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria, 2025. Acessado em: 25 de julho de 2025. Disponível em: <https://www.R-project.org/>.

WICKHAM, H. **forcats**: Tools for Working with Categorical Variables (Factors). R package version 1.0.0, 2023. Acessado em: 25 de julho de 2025. Disponível em: <https://forcats.tidyverse.org/>.

WICKHAM, H.; FRANÇOIS, R.; HENRY, L.; MÜLLER, K.; VAUGHAN, D. **dplyr**: A Grammar of Data Manipulation. R package version 1.1.4, 2025. Acessado em: 25 de julho de 2025. Disponível em: <https://dplyr.tidyverse.org>.

WICKHAM, H.; GROLEMUND, G. **R for Data Science**: Import, Tidy, Transform, Visualize, and Model Data. O'Reilly Media, 2017.