

Alavancando o poder do RMarkdown com as linguagens da Web e D3.js para produzir histórias de dados envolventes sobre Finanças Públicas

Anônimo

Palavras chave: R Markdown - Data Visualization - Data Storytelling - Web Standards - r2d3 - Public Finance - Government

URL da Aplicação: <https://tchiluanda.github.io/DC/>

Introdução

O R Markdown permite integrar a análise e a comunicação de dados, costurando texto e códigos numa miríade de formatos de saída, desde artigos e slides até livros e aplicações web interativa.

A simplicidade da linguagem markdown, o poder do R e a versatilidade de formatos de saída fazem com que usuários de R possam produzir, num mesmo ambiente, um fluxo completo e reproduzível, desde a importação de dados brutos até a construção de um documento visualmente atraente (como ilustra Allison Horst em: https://github.com/allisonhorst/stats-illustrations/blob/master/rstats-artwork/rmarkdown_wizards.png).

Em particular, o formato de saída HTML permite que usuários sem conhecimentos de web design produzam páginas web completas, com design agradável e prontas para publicação e comunicação do resultado de suas análises.

Indo além

No entanto, é possível criar histórias de dados ainda mais envolventes e atraentes mergulhando um pouco mais fundo nas linguagens e tecnologias envolvidas na geração de documentos HTML para a Web a partir do R Markdown.

Uma das características do R Markdown que lhe conferem tamanha versatilidade é a possibilidade de inclusão de “chunks” de códigos escritos em outras linguagens além de R.

Por exemplo, para personalizar um pouco mais a aparência da página web, é fundamental usar CSS (“Cascading Style Sheets”), a linguagem usada para descrever a apresentação de páginas Web, incluindo cores, layouts e fontes. Códigos em CSS podem ser incluídos no próprio arquivo R Markdown. No documento HTML gerado, esses chunks serão incluídos como blocos `<style>`.

Para incluir novos elementos visuais que possam ser formatados pelo código em CSS (boxes explicativos em meio ao texto, como no caso do nosso projeto), é possível utilizar a própria sintaxe do Pandoc, em muitos casos, como no caso de “fenced divs” e “bracketed spans”. Para os demais casos, existe a possibilidade de se incluir o código HTML puro no corpo do texto.

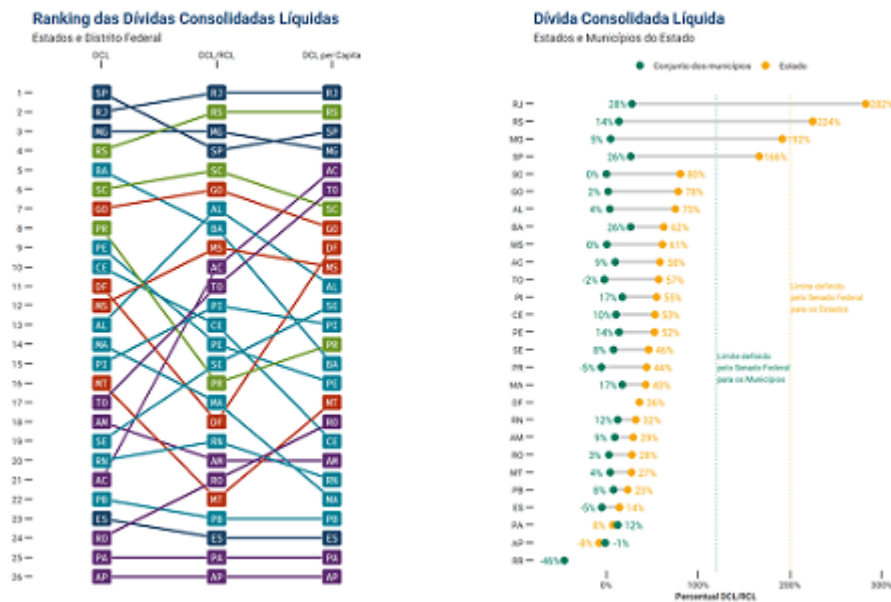
Além disso, é possível acrescentar interatividade à página com o uso de Javascript, cujos códigos também podem ser incluídos como chunks. Por fim, graças ao pacote `r2d3`, pode-se ainda incluir chunks de D3.js, a principal biblioteca em Javascript para construção de visualizações interativas e sem formatos pré-determinados.

O Projeto

Neste projeto, utilizamos as demonstrações financeiras declaradas ao Tesouro Nacional pelo governo federal, pelos governos estaduais e por 93% dos 5.570 municípios brasileiros. Com base nesses dados, construímos com

R Markdown uma página Web apresentando as informações das dívidas públicas dos governos brasileiros numa estrutura narrativa e envolvente. Para isso utilizamos diversos recursos.

Toda a importação, preparação, processamento, agrupamento e tratamento dos dados foi feita com R, utilizando extensivamente os pacotes do tidyverse. Para os gráficos estáticos, utilizamos ggplot2.



Para encontrar um equilíbrio entre simplicidade e rigor conceitual no texto, deixamos o texto principal mais simples, mas incluímos a possibilidade de o usuário clicar em termos mais técnicos para que a página exiba quadros com explicações adicionais.

Isso é possível com uma combinação do uso da sintaxe do Pandoc para criar elementos customizados, chunks de CSS (para a formatação dos elementos) e Javascript (para acrescentar interatividade).

Além disso, como mencionado e indicado acima, como experimento fizemos uso do pacote r2d3 para construir gráficos em D3 diretamente no R Markdown. Uma vantagem de usar essa abordagem é permitir a utilização de gráficos interativos no próprio formato HTML, sem a necessidade de um servidor Shiny.

Finalmente, utilizamos o chunk de CSS para definir o layout geral da página, fontes, cores, espaçamentos, margens etc.

Comunicando Finanças Públicas de uma forma mais amigável para a sociedade

Assim, para poder combinar “Finanças Públicas” com “histórias de dados envolventes”, ou seja, para tornar esse assunto menos árido e mais interessante, decidimos fazer uso do potencial das linguagens dos padrões da Web, reunidas no ambiente ao qual já estamos tão bem familiarizados para realizar nossas análises de dados.

Dessa forma, é possível, num único documento de texto simples .Rmd, escrever textos, importar e manipular dados com R, gerar visualizações complexas e atraentes com ggplot2, construir elementos visuais com markdown e HTML, formatar a aparência desses elementos com CSS, incluir interatividade com Javascript, embutir visualizações interativas e completamente customizadas com D3.

Além da versatilidade, a vantagem do R Markdown é a de integrar todo o fluxo desde a importação dos dados originais até a criação da página Web final. Assim, nas atualizações futuras dos dados, basta “costurar” novamente o arquivo .Rmd para se obter a página Web atualizada.

Em nossa apresentação, pretendemos mostrar os detalhes da implementação e compartilhar aprendizados e dicas que podem inspirar e contribuir para a construção de história de dados visualmente atraentes com R Markdown. O código está disponível em: <https://github.com/tchiluanda/DC>, e a aplicação está on-line em: <https://tchiluanda.github.io/DC/>.