

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

Tiago Pereira , Fernando Barbalho , Jordao Goncalves , Lucas Leite

Palabras clave: R Markdown - Data Visualization - Data Storytelling - Web Standards - r2d3 - Public Finance - Government

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.

A figura abaixo, desenhada por Allison Horst, é uma excelente ilustração desse processo:



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” (*Pandoc User Guide*, n.d.). 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 (*Data-Driven Documents*, n.d.), a principal biblioteca em Javascript para construção de visualizações interativas e sem formatos pré-determinados.

O Projeto: comunicando Finanças Públicas de uma forma mais amigável para a sociedade

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 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.

Antes disso, porém, é preciso garantir que estamos falando de coisas iguais. No caso da dívida consolidada, entram, essencialmente, a dívida mobiliária, quando o país emite títulos públicos para captar dinheiro; a dívida contratual, quando os governos estaduais e municipais recorrem a empréstimos e (re)financiamentos, inclusive com o próprio governo federal; e a dívida de precatórios que venceram e não foram pagos, aquelas impostas por ações judiciais.



Antes disso, porém, é preciso garantir que estamos falando de coisas iguais. No caso da dívida consolidada, entram, essencialmente, a dívida mobiliária, quando o país emite títulos públicos para captar dinheiro; a dívida contratual, quando os governos estaduais e municipais recorrem a empréstimos e (re)financiamentos, inclusive com o próprio governo federal; e a dívida de precatórios que venceram e não foram pagos, aquelas impostas por ações judiciais.

Dívida Consolidada e as várias formas de se medir a dívida pública

Na verdade, existem diversas formas de se medir a dívida de um ente público. É possível que você já tenha esbarrado em alguns desses: DPF, DPMF, DBOG, DLSB, DC, uma verdadeira sopa de letrinhas.

A Dívida Consolidada ou Fundada* corresponde ao montante total das obrigações financeiras, apurado sem duplicidade (excluídas obrigações entre órgãos da administração direta e entre estes e as entidades da administração indireta), assumidas:

- pela emissão de títulos públicos, para amortização em prazo superior a 12 (doze) meses (dívida mobiliária);
- em virtude de leis, contratos (dívida contratual), convênios ou tratados e da realização de operações de crédito*, para amortização em prazo superior a 12 (doze) meses;

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).

```

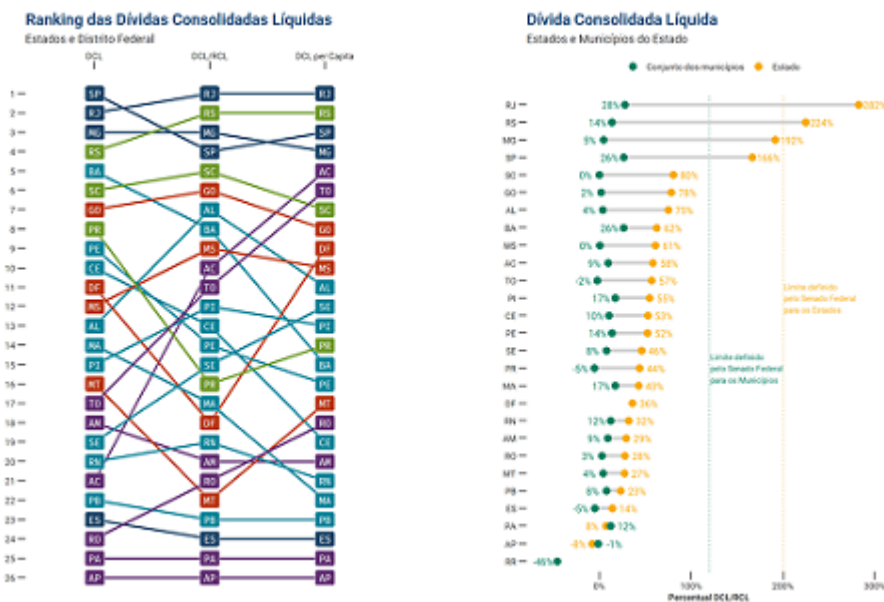
16 // (cs, echo=FALSE)
17
18 #-----
19 padding: 2px;
20 background-color: lightblue;
21 cursor: pointer;
22 }
23
24 .descricao {
25   background-color: lightblue;
26   margin: 5px;
27   font-size: 0.7em;
28   font-family: "Roboto";
29 }
30
3132
3133 // (cs, echo=FALSE)
3134 const conceitos = document.getElementsByClassName('conceito');
3135 const descricoes = document.getElementsByClassName('descricao');
3136
3137 let clicados = {};
3138
3139 for (let i = 0; i < descricoes.length; i++) {
3140   descricoes[i].style.display = 'none';
3141 }
3142
3143 document.addEventListener("click", function (event) {
3144
3145   if (event.target.matches('.conceito')) {
3146     let idConceito = event.target.id;
3147     console.log("clicou em", idConceito);
3148
3149     identificador = idConceito.substring(idConceito.indexOf("-")+1);
3150
3151 Antes disso, porém, é preciso garantir que estamos falando de coisas iguais. No caso da dívida consolidada, conceito=consolidado-cc, então, essencialmente, a dívida mostramos, quando se trata de títulos públicos para conter empréstimo a dívida contratada quando os governos estaduais e municipais recorrem a empréstimos e re/financiamentos, inclusive com o priorio governo federal e a dívida de [proteções], conceito Empréstimo/prestáculos que vencerão e não foram pagos, aquelas impostas por ações judiciais.
3152
3153 //--- Dívida Consolidada e as várias formas de se medir a dívida pública
3154 Na verdade, existem diversas formas de se medir a dívida de um ente público. É possível que você já tenha esbarhado em alguns desses: SPF, GDFM, DGF, CLDP, DC... uma verdadeira sopa de letrinhas.
3155
3156 // (d) Parâmetros CC
3157 const dataset = datasettable, variavel) {
3158   const valores_unicos = dataset
3159     .map(c => d[variavel])
3160     .filter((v, i, a) => a.indexOf(v) === i);
3161   return valores_unicos
3162 }
3163
3164 // popular os menus
3165 const $menu_estado = d3.select("#menu-estado");
3166 const $menu_municipio = d3.select("#menu-municipio");
3167
3168 let $columnas_estados = Object.keys(data["estados"]);
3169 //let $columnas_municipios = Object.keys(data["municipios"]);
3170
3171 let estados = data["estados"];
3172 .map(d => ({ nome_estado: d["nome"],
3173             "uf": d["UF"],
3174             "DC": d["DC"],
3175             "DCL": d["DCL"]
3176           })

```

Chunks de CSS, Javascript e D3.js num mesmo documento R Markdown.

No alto, à direita: sintaxe do Pandoc para criar `` e `<div>`.

Para os gráficos estáticos, utilizamos ggplot2 (exemplos abaixo).



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. Gráficos bonitos feitos com o `ggplot2` auxiliam bastante na tentativa de tornar o conteúdo mais interessante. Mas investir um pouco de tempo para conhecer as possibilidades de CSS, principalmente, pode trazer excelentes retornos para a qualidade final do produto.

Reflexões

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 obtermos a página Web atualizada.

A apresentação de Desirée de Leon (*Of Teacups, Giraffes, & R Markdown* 2020) na RStudio::conf 2020, serviu de grande inspiração para este trabalho. Como Desirée diz, na apresentação: “a preocupação com o design não apenas causa uma melhor impressão inicial, mas pode melhorar a legibilidade, a navegação e a experiência geral do usuário à medida em que ele consome o conteúdo”.

O código está disponível em: <https://github.com/tchiluanda/DC>

A aplicação está on-line em: <https://tchiluanda.github.io/DC/>

Referências

Data-Driven Documents. n.d. <https://d3js.org/>.

Of Teacups, Giraffes, & R Markdown. 2020. <https://rstudio.com/resources/rstudioconf-2020/of-teacups-giraffes-r-markdown/>.

Pandoc User Guide. n.d. <https://www.pandoc.org/MANUAL.html#divs-and-spans>.

Tiago Pereira
National Treasury of Brazil
tiagombp@gmail.com

Fernando Barbalho
National Treasury of Brazil
fernando.barbalho@tesouro.gov.br

Jordao Goncalves
National Treasury of Brazil
jordao.m.goncalves@tesouro.gov.br

Lucas Leite
National Treasury of Brazil
lucas.leite@tesouro.gov.br