

14º Congresso Brasileiro de Design: Visualizando a COVID-19 associada a outras variáveis: uma análise comparativa de ferramentas interativas

14th Brazilian Congress on Design Research: Visualizing COVID-19 related to other variables: a comparative analysis of interactive tools

FALCONIERI, Igor; Mestrando; Universidade Federal do Rio de Janeiro

igorfalconieri@gmail.com

KOSMINSKY, Dóris; Doutora; Universidade Federal do Rio de Janeiro

doriskos@eba.ufrj.br

Desde o início da pandemia da COVID-19, técnicas e ferramentas de visualização de dados foram aplicadas no suporte ao entendimento da doença e à tomada de decisão para contê-la. Parte desse entendimento passa por verificar como outras variáveis intensificam os impactos da COVID-19 e são intensificadas por ela. Esse artigo se dedica à análise comparativa de uma seleção de ferramentas interativas de visualização de dados disponíveis na web que possibilitam a observação de relações entre eventos de COVID-19 e outras variáveis. O objetivo desta análise é reconhecer nessas ferramentas padrões de design, de forma a identificar critérios prescritivos de concepção e projeto. Foram analisadas 33 ferramentas, agrupadas por dois formatos de apresentação (*dashboards* e artigos de *scrollytelling*) e classificadas por características de dois grupos (dados/domínio e técnicas). Oito padrões que representam escolhas de design foram observados e discutidos no artigo.

Palavras-chave: Visualização de dados; Interfaces digitais; COVID-19.

*Since the beginning of the COVID-19 pandemic, data visualization techniques and tools support understanding of the disease and the decision-making to contain it. Part of this understanding and decision-making involves verifying how other variables intensify the impacts of COVID-19 and are intensified by it. This article presents a comparative analysis of selected interactive data visualization web tools that allow the observation of relationships between COVID-19 events and other variables. The analysis intends to recognize patterns and regularities in these tools, identifying prescriptive criteria for design. Thirty-three tools were analyzed, grouped by two presentation formats (*dashboards* and *scrollytelling* articles), and classified by characteristics of two groups (*data/domain* and *techniques*). The article notes and discusses eight patterns representing design choices.*

Keywords: Data visualization; Digital interfaces; COVID-19.

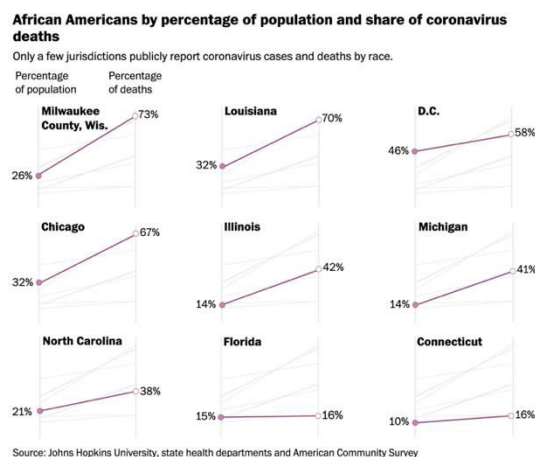
1 Introdução

Desde que se identificou o surto da Doença por Coronavírus 2019 (COVID-19), e especialmente depois dele ser reconhecido como pandemia, em março de 2020, o mundo se dedicou em grande esforço ao entendimento de como essa doença se espalha e como é possível contê-la. Por ocorrer num momento de extensa produção e compartilhamento de dados, observa-se nesse espaço de tempo uma proliferação de visualizações a respeito da COVID-19 (GIANNELLA; VELHO, 2021). O professor de jornalismo visual da Universidade de Miami Alberto Cairo expõe que “os gráficos moldaram nossas percepções da pandemia” e classifica essa como “a primeira crise global visualizada na história” (2020, pp. 195 e 196). Técnicas e ferramentas de visualização de dados foram aplicadas em grande escala no suporte ao entendimento da doença e à tomada de decisão em medidas de contenção.

Parte desse entendimento e tomada de decisão passa também por verificar como outras variáveis intensificam os impactos da COVID-19 e também são intensificadas por ela. Em diversos casos, essas variáveis podem se configurar como condições de **vulnerabilidade**, fornecendo hipóteses do porquê os impactos diretos e indiretos da doença são desiguais entre diferentes grupos. Fatores sanitários, como a prevalência de comorbidades e a diferença de acesso a planos de saúde, públicos ou privados (LI et al., 2021); e socioeconômicos, como o nível de renda, raça, local de residência, condições de domicílio, escolaridade e possibilidade de trabalho remoto (PATEL et al., 2020) (LI et al., 2021), explicam diferenças em eventos de contaminação, hospitalização, morte e vacinação na distribuição populacional.

Tão logo após a eclosão da pandemia, ainda em abril de 2020, o jornal Washington Post publicou uma análise com recursos visuais (Figura 1) que chamava atenção para a alarmante e desproporcional taxa de mortalidade por COVID-19 entre afro-americanos (THEBAULT; TRAN; WILLIAMS, 2020). No Brasil, Li, et al (2021) e Ferreira (2020) encontraram no estado e na cidade de São Paulo, respectivamente, maior chance de infecção, hospitalização e morte por COVID-19 entre pacientes que vivem nas áreas mais pobres, quando comparado àqueles que vivem nas áreas mais ricas.

Figura 1 - Gráfico slope do artigo *The coronavirus is infecting and killing black Americans at an alarmingly high rate*



Fonte: The Washington Post, 2020.

Esses não são os únicos exemplos de análises que levaram em conta possíveis fatores de vulnerabilidade para acompanhar a evolução da COVID-19. De fato, uma diversidade de ferramentas de visualização abordou outras variáveis associadas à doença. Explorar a concomitância desses eventos, no entanto, impõe desafios para a visualização de dados: a necessidade de permitir na visualização o entendimento de relações estatísticas torna mais complexa a codificação visual. É preciso empregar parâmetros de representação a fim de que, em uma ferramenta de visualização, eles possam “substituir cálculos cognitivos por inferências perceptivas simples e melhorar a compreensão, a memória e a tomada de decisões” (HEER; BOSTOCK; OGIEVETSKY, 2010).

Esse artigo se dedica à análise comparativa de uma seleção de ferramentas que possibilitam a observação de relações entre eventos de COVID-19 e outras variáveis. O objetivo dessa análise é reconhecer nessas ferramentas padrões e regularidades de design, identificando critérios prescritivos de concepção e projeto. A seguir serão descritos os critérios de seleção, agrupamento e classificação e comentadas as descobertas obtidas.

2 Metodologia

2.1 Critérios de seleção

As ferramentas de visualização analisadas foram eleitas por quatro critérios e obtidas por dois métodos de busca. O primeiro deles foi uma pesquisa em sites de (1) empresas e designers especializados em visualização de dados; (2) *think tanks* voltados para políticas públicas; (3) órgãos de saúde; e (4) investidores na área de saúde e políticas públicas. A tabela 3, no [apêndice](#), especifica por ordem alfabética todos os sites buscados de cada agrupamento.

O segundo método de busca utilizou a ferramenta COVIC: COVID-19 Online Visualization Collection, um repositório de visualizações da COVID-19 publicadas na internet mantido pela Northeastern University, dos Estados Unidos. O arquivo agrupa mais de 3 mil artigos contendo visualizações que são classificadas através de metadados (CIUCCARELLI; KAHN, 2021). Três desses metadados permitiram a escolha de filtros de busca voltados para resultados que melhor atendessem os critérios de elegibilidade para a análise: *language* (idioma), *subject* (assunto) e *article technique* (técnica do artigo). A tabela 4, no [apêndice](#), detalha os filtros aplicados em cada um dos três metadados. Em *language* foram filtrados os artigos em inglês e português, por serem de domínio fluente do autor; em *subject* foram selecionados os filtros *economic* (econômicos), *environmental* (ambientais) e *social* (sociais); e em *article technique* foram filtrados os formatos *dashboard* e *scrollytelling*, que empregam técnicas de interação e que serão explicados a seguir.

A seleção dos resultados de busca atendeu aos seguintes critérios de elegibilidade: (1) disponibilidade das ferramentas na web; (2) emprego de técnicas de interação na visualização; (3) apresentação de variáveis próprias da COVID-19; e (4) apresentação de variáveis preexistentes à COVID-19, ou seja, que não são próprias ou não passaram a ser medidas por causa da doença. As variáveis dos critérios 3 e 4 estão especificadas por ordem alfabética nas tabelas 5 e 6, respectivamente, no [apêndice](#). Elas também servirão como critérios de classificação, apontando quais das ferramentas analisadas as exibe. Os métodos de classificação, bem como o agrupamento das variáveis, serão descritos mais adiante no item 2.3.

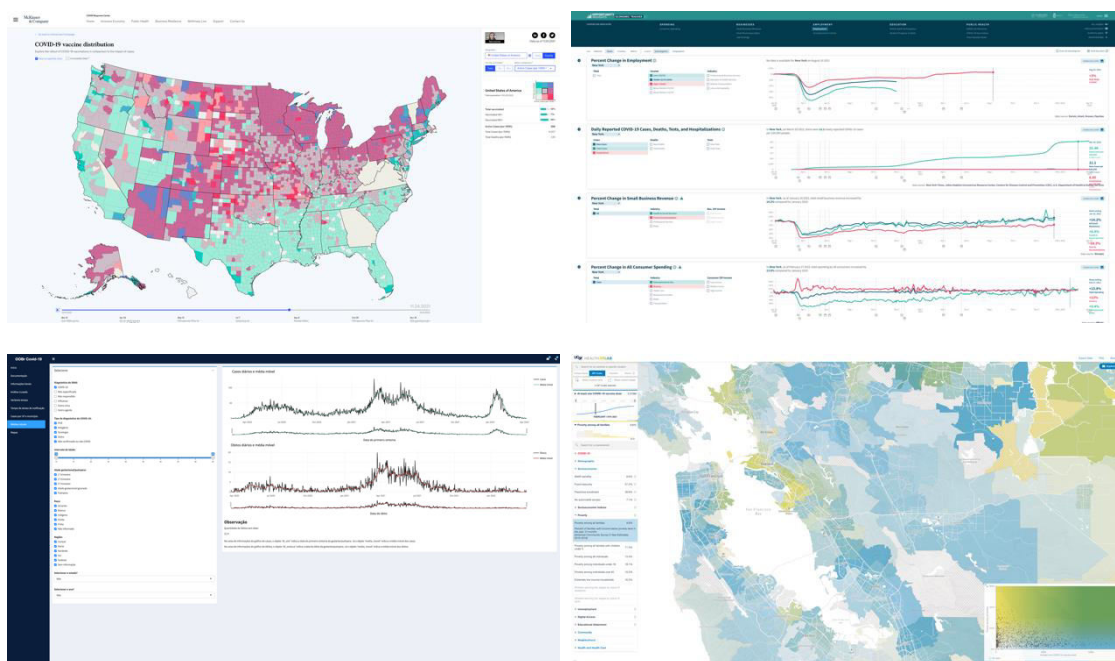
2.2 Critérios de agrupamento

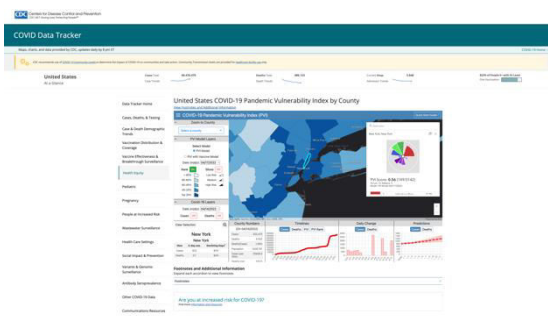
Os dois métodos de busca forneceram, no conjunto, 33 ferramentas de visualização que atenderam a todos os critérios de elegibilidade. As ferramentas analisadas foram agrupadas por dois formatos de apresentação: *dashboard* e *scrollytelling*.

Define-se como *dashboard* “uma exibição visual das informações mais importantes necessárias para atingir um ou mais objetivos; consolidada e organizada em uma única tela de modo que as informações possam ser monitoradas rapidamente” (FEW, 2006). Essa definição de Few salienta que *dashboards* apresentam visualmente a informação com ênfase em gráficos, mesmo que textos sirvam como suporte; e permitem que uma coleção de informações comumente vindas de diversas fontes possa ser sumarizada numa leitura “de relance”. Além disso, reforça que, em *dashboards*, a informação deve “caber em uma única tela, inteiramente disponível dentro do alcance do olho do leitor” (FEW, 2006). Essa descrição faz com que algumas das ferramentas analisadas sejam definidas por Few como *multiple dashboards*, por terem mais de uma página, normalmente variando subdomínios de um mesmo assunto. Para fins práticos de análise, ferramentas com múltiplas páginas foram tratadas nas tabelas de análise (1 e 2) de forma agrupada pelo título da coleção.

Ao todo 18 ferramentas analisadas na amostra têm apresentação em formato de *dashboard*. Alguns exemplos são as ferramentas *COVID-19 vaccine distribution*, da McKinsey (Figura 2); *Opportunity Insights Economic Tracker*, das universidades Harvard e Brown em parceria com a Fundação Bill & Melinda Gates (Figura 3); *Observatório Obstétrico Brasileiro COVID-19*, da Plataforma de Ciência de Dados aplicada à Saúde da Fiocruz com apoio do estúdio ODD (Figura 4); *UCSF Health Atlas*, da Universidade da Califórnia em São Francisco (Figura 5); *Belgium COVID-19 Epidemiological Situation*, da Sciensano (Figura 6); e *COVID Data Tracker*, do Centers for Disease Control and Prevention, o CDC (Figura 7).

Figuras 2 a 7: Capturas de tela de exemplos de *dashboards* analisados na amostra.





Fonte: Figura 2, McKinsey; Figura 3, Harvard University/Brown University/Bill & Melinda Gates Foundation; Figura 4, PCDS/ODD Studio; Figura 5, University of California, San Francisco; Figura 6, Sciensano; Figura 7, Centers for Disease Control and Prevention.

Scrollytelling, por sua vez, é um nome do segmento jornalístico, e descreve artigos digitais de longo formato que normalmente incorporam recursos multimídia para narrar histórias não-ficcionais em detalhes; uma combinação dos termos *storytelling* e *scrolling* (SEYSER; ZEILLER, 2018). O primeiro termo é o ato de narrar histórias e o segundo o de rolar uma página da web. Esse formato de artigo incorpora recursos de visualização de dados como um suporte narrativo:

“Como os artigos de *scrollytelling* geralmente incluem muitas imagens e elementos multimídia e são usados para contar histórias complexas, infográficos podem ser encontrados com frequência nesses artigos” (SEYSER; ZEILLER, 2018, p. 402).

Das ferramentas da amostra, 15 apresentam os dados no formato de *scrollytelling*. Alguns exemplos são os artigos *More Than 1.08 Billion Shots Given: Covid-19 Tracker*, da Bloomberg (Figura 8); *COVID-19 Vaccination Tracker*, da Reuters (Figura 9); *1 in 5 Prisoners in the U.S. Has Had COVID-19*, do The Marshall Project (Figura 10); *See How Rich Countries Got to the Front of the Vaccine Line* (Figura 11) e *100 New Yorkers* (Figura 12), ambos do The New York Times; e *COVID-19 no Brasil: Marcas da Pandemia*, da Valor Econômico (Figura 13).

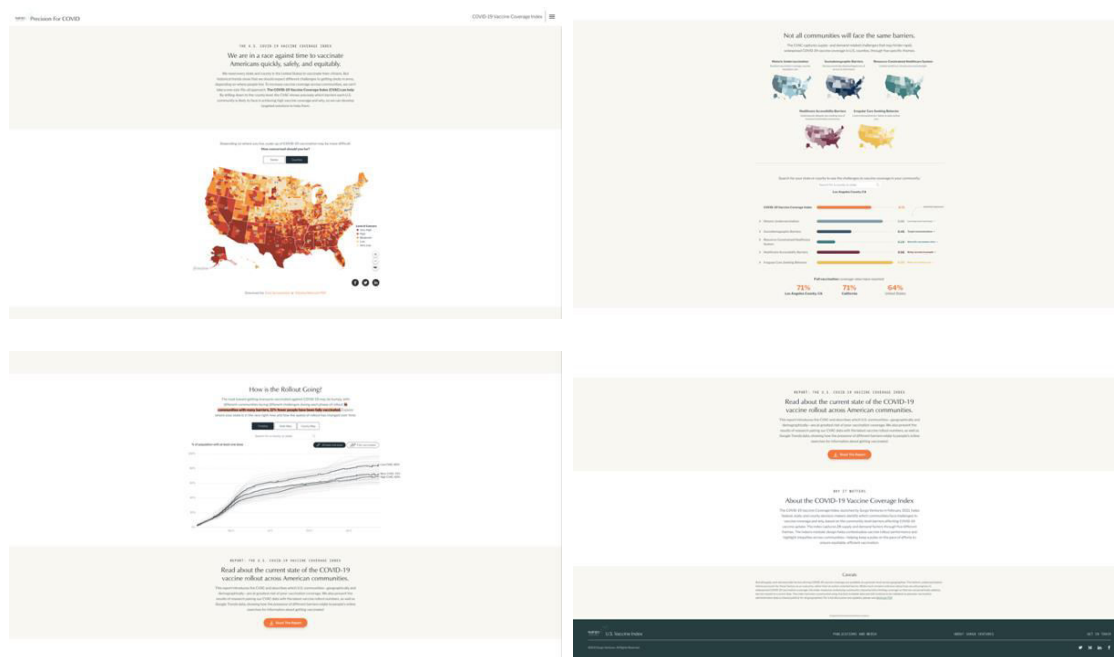
Figuras 8 a 13: Capturas de tela de exemplos de *scrollytelling* analisados na amostra.



Fonte: Figura 8, Bloomberg; Figura 9, Reuters; Figura 10, The Marshall Project; Figuras 11 e 12, The New York Times; Figura 13, Valor Econômico.

Alguns *dashboards* e artigos de *scrollytelling*, no entanto, podem se parecer. Um exemplo interessante é o *dashboard The U.S. Covid-19 Vaccine Coverage Index*, da Surgo Ventures (Figuras 14 a 17), que tem a disposição da informação semelhante à que é normalmente encontrada em artigos de *scrollytelling* – sequencial e vertical. Um fator de distinção, nesse caso, é que essa ferramenta não é focada no recorte de uma história, nem desenvolvida em torno de um texto narrativo (SEYSER; ZEILLER, 2018). Seu foco é exibir, ainda que numa disposição visual diferente, uma combinação de informações aberta a múltiplas explorações, função característica dos *dashboards*.

Figuras 14 a 17: Capturas de tela exibindo a sequência de rolagem da ferramenta *The U.S. Covid-19 Vaccine Coverage Index*.



Fonte: Surgo Ventures, 2021.

2.3 Critérios de classificação

As ferramentas nos dois formatos de apresentação foram analisadas uma a uma e classificadas por características de dois grandes grupos: o dos **dados/domínio** e o das **técnicas**. O primeiro grande grupo, dados/domínio, compreende as características (1) tipos de dado e medidas, especificados na tabela 7, no [apêndice](#); (2) variáveis próprias da COVID-19, descritas na tabela 5, no [apêndice](#); (3) variáveis preexistentes, descritas na tabela 6; no [apêndice](#) (4) disponibilidade da metodologia; e (5) disponibilidade de download dos dados.

O segundo grupo, técnicas, abarca as características (1) função de visualização; (2) estrutura de visualização, sendo essa e a anterior baseadas na taxonomia do estúdio dinamarquês Ferdio, *Data Viz Project* (2017), e sumarizadas nas tabelas 8 e 9, no [apêndice](#); (3) técnicas de interação, divididas pelas taxonomias de Yi, *et al* (2007) e de Shneiderman (1996) e especificadas nas tabelas 10 e 11; e (4) tecnologia, agregando linguagens de programação, empresas desenvolvedoras e plataformas de *GIS* e *business intelligence* usados para prover a ferramenta, quando informados.

A identificação das características de dados/domínio e de técnicas em cada uma das ferramentas de visualização está exibida nas tabelas 1 e 2, respectivamente. Cada uma delas foi repartida em A e B para facilitar a leitura nesse artigo. Além disso, todos os nomes nas tabelas de análise foram mantidos em inglês para fins de padronização e por facilidade de aderência às taxonomias utilizadas. No [apêndice](#) desse artigo, é possível encontrar a tradução para o português de todas as variáveis, dos tipos de dados e medidas e de alguns exemplos de estrutura de visualização. A tabela de análise completa, sem repartições, bem como os links de todas as ferramentas analisadas, podem ser acessados clicando [aqui](#).



14º Congresso Brasileiro de Design
ESDI Escola Superior de Desenho Industrial
ESPM Escola Superior de Propaganda e Marketing

Tabela 1A – Análise de ferramentas por características dos dados/domínios abordados

Displaying Format				Data Types/Measures		Data/Domain																				#	Methodology Available	Data Download Available	
	Source	Country	Name	Measures	Data Types	COVID-Caused Variables																							
						Quantity	Score	Time	Variation	Rate	Level	Cases/Infections	CLU/ILI	Contact tracing	Deaths	Hospital admissions	Hospital usage/availability	ICU/ventilators usage/availability	Mobility/Social distancing	Positivity/Seroprevalence	Risk of death	Risk of infection	Stringency	Symptoms	Testing				Transmission
Dashboard	Arizona State		Arizona Department of Health Services COVID-19 Data																									11	
	CDC		Centers for Disease Control and Prevention COVID Data Tracker																									13	
	Columbia U.		Columbia University COVID-19 Risk Mapping																									4	
	Cuebiq		Social Inequality Analysis																									1	
	Emory U.		Emory University COVID-19 Health Equity Interactive Dashboard																									5	
	ESRI		Impact Planning for COVID-19																									4	
	Harvard U./Brown U.		Opportunity Insights Economic Tracker																									7	
	IBICT/MCTI		Visão COVID-19																									7	
	McKinsey		COVID-19 vaccine distribution																									3	
	Milwaukee County		Milwaukee County COVID-19 Surveillance Dashboard																									8	
	PCDaS/ODD Studio		Observatório Obstétrico Brasileiro COVID-19																									7	
	Rainforest Foundation US		Map to Optimize COVID-19 Response in Panama's Indigenous Communities																									1	
	Sciensano		Belgium COVID-19 Epidemiological Situation																									10	
	Surgo Ventures		Precision for COVID: COVID-19 Vaccine Coverage Index																									5	
	Surgo Ventures		Precision for COVID: US Data Explorer																									5	
	U. of California SF		UCSF Health Atlas																									6	
	U. of Maryland		COVID-19 Impact Analysis Platform																									8	
	Washington State		Washington State Department of Health COVID-19 Data Dashboard																									12	
Scrolling	Bloomberg		More Than 1.08 Billion Shots Given: Covid-19 Tracker																									4	
	Bloomberg		Half of U.S. States Have Big Racial Vaccine Gaps Heading Into Omicron																									1	
	Reuters		COVID-19 Vaccination Tracker																									4	
	Reuters		The social distancing of America																									1	
	The Marshall Project		1 in 5 Prisoners in the U.S. Has Had COVID-19																									1	
	The Marshall Project		A State-By-State Look at 15 Months of Coronavirus in Prisons																									2	
	The New York Times		100 New Yorkers																									1	
	The New York Times		Location Data Says It All: Staying at Home During Coronavirus Is a Luxury																									1	
	The New York Times		One-Third of All U.S. Coronavirus Deaths Are Nursing Home Residents or Workers																									2	
	The New York Times		Pandemic's Racial Disparities Persist in Vaccine Rollout																									1	
	The New York Times		See How Rich Countries Got to the Front of the Vaccine Line																									2	
	The New York Times		What Are the Vaccine Roadblocks Where You Live?																									1	
	The Washington Post		How a surge of coronavirus patients could stretch hospital resources in your area																									4	
	Valor Econômico		COVID-19 no Brasil: Marcas da Pandemia																									2	
	Vox		US Covid-19 deaths, explained in 8 charts and maps																									2	

Obs.: As colunas intituladas # exibem a soma de variáveis diferentes encontradas na ferramenta.

Fonte: o autor.

Data/Domain			
Preexisting Variables			
Socioeconomical	Demographical		
Educational	Environmental		
Health factors			
Assistance/Welfare	Birth	Addictions	Comorbidities
Business activities	Critical risk environments	Critical risk workers	Critical risk environments
Compulsory settings	Disability	Disability	Disability
Confinement	Health professionals availability	Health professionals availability	Health professionals availability
GDP/GDP per capita	Healthcare system	Healthcare system	Healthcare system
Government approval	Infectious diseases (not COVID)	Infectious diseases (not COVID)	Infectious diseases (not COVID)
Government spending	Insurance access	Insurance access	Insurance access
Hunger/Food insecurity	Life expectancy	Life expectancy	Life expectancy
Incarceration	Mental health	Mental health	Mental health
Income	Non-communicable diseases	Non-communicable diseases	Non-communicable diseases
Inequality	Nutrition/Obesity	Nutrition/Obesity	Nutrition/Obesity
Infraction	Physical inactivity	Physical inactivity	Physical inactivity
Labour	Pregnancy	Pregnancy	Pregnancy
Political affiliation	Preventive care	Preventive care	Preventive care
Poverty	Screening	Screening	Screening
Production	Testing access	Testing access	Testing access
Social vulnerability	Vaccination (not COVID)	Vaccination (not COVID)	Vaccination (not COVID)
Socioeconomic status			
Stock market indexes			
Violence			
Age			
Density			
Foreign born			
Household composition			
Housing			
Minorities			
Population			
Race/Ethnicity			
Rent/Eviction			
Sex/Gender			
Transportation			
Urban/rural			
Access to information			
Childhood poverty			
Computer access			
Educational attainment			
Internet access			
Limited language proficiency			
Student participation			
Student progress			
Air quality			
Bad weather risk			
Drinking water			
Heat/cold days			
Pesticides			
Pollution			
Sewage			

Fonte: o autor.

Tabela 2A – Análise de ferramentas por características das técnicas empregadas

[illegible]

Obs. 1: Na seção de estruturas de visualização (*Visualization structures*), algumas estruturas são classificadas em duas famílias ao mesmo tempo. Elas são repetidas na tabela e identificadas com a cor correspondente à segunda família próximo do nome.

Obs. 2: As colunas intituladas # exibem a soma de técnicas diferentes empregadas na ferramenta. Quando elas são repetidas na tabela, contam apenas uma vez.

Fonte: o autor.

[illegible]

Obs. 2: As colunas intituladas # exibem a soma de técnicas diferentes empregadas na ferramenta. Quando elas são repetidas na tabela, contam apenas uma vez.

Fonte: o autor.

3 Observações de design

Na análise da amostra, foram observados oito padrões que estão numerados e destacados em vermelho nas tabelas 1 e 2. Esses padrões buscam gerar aprendizado sobre escolhas de design na concepção e desenvolvimento das ferramentas e prover elementos prescritivos para um projeto prático futuro.

O primeiro padrão (Tabelas 1A e 1B (1)) aponta para a diferença de abundância e diversidade de dados entre os dois grupos de formato de apresentação, *dashboard* e *scrollytelling*. Os *dashboards* exibem significativamente muito mais variáveis, tanto próprias da COVID-19 quanto preexistentes. Diversos deles passam das dezenas de variáveis, enquanto alguns artigos de *scrollytelling* aparecem até mesmo com apenas uma. Essa diferença ocorre pela natureza com que cada formato apresenta a informação: *dashboards* para finalidades analíticas devem permitir exploração para o exame de dados complexos e a descoberta de relações entre eles (FEW, 2006), porém, o caminho dessa exploração é muito mais aberto se comparado ao formato *scrollytelling*. Os artigos de *scrollytelling* são frequentemente centrados no texto e numa estrutura de apresentação narrativa (SEYSER; ZEILLER, 2018), o que limita a exploração às variáveis que dão suporte à história. Nos *dashboards*, por outro lado, é a finalidade de descoberta que pode justamente ter orientado a escolha desse formato de apresentação; ou, como descreve Munzner (2014, p. 47), “a motivação para linguagens interativas sofisticadas, pois o designer de visualização não sabe de antemão o que o usuário precisará ver”.

O segundo padrão (tabela 1B (2)) sugere uma subutilização de variáveis ambientais, encontradas somente em *dashboards* bastante ricos em variáveis e de múltiplas páginas. O terceiro (tabela 2A (3)) destaca o uso frequente de soluções mais familiares em estruturas de visualização de dados: gráficos de barras (*bar chart*), barras de progresso (*progress bar*), tabelas (*table chart*) e gráficos de linha (*line graph*) são encontrados predominantemente ao longo de quase todas as ferramentas. No mesmo sentido, o quarto padrão (tabela 2A (4)) salienta o uso do mapa coroplético (*choropleth map*) como estrutura preferencial para a apresentação de dados geoespaciais.

Uma hipótese que pode explicar os dois últimos padrões está na escolha de design em favor da efetividade e da facilidade de reconhecimento por parte do usuário. Posição e comprimento são os dois canais de maior acurácia de interpretação para a percepção humana segundo os estudos clássicos de Cleveland e McGill, de 1984, reiterados pelos estudos de Heer e Bostock, de 2010 (MUNZNER, 2014, pp. 104 e 105). Ambos são canais que expressam magnitude em gráficos de barra, barras de progresso e gráficos de linha, o que torna essas estruturas de visualização seguras na intenção de garantir efetividade entre as escolhas de design. Para além disso, o fator da familiaridade, ou da facilidade de reconhecimento e leitura, pode estar relacionado à impossibilidade de estimar o nível de letramento em visualização de dados (DVL, *data visualization literacy*) de todos os possíveis usuários. Börner, *et al* (2016) definem DVL como “a capacidade de interpretar e de obter significado a partir de padrões, tendências e correlações em representações visuais de dados.”

No domínio do letramento em visualização de dados, Börner, *et al* encontraram em estudos com visitantes de três museus de ciências nos Estados Unidos uma alta precisão de reconhecimento de gráficos de barras e gráficos de linha comparados a outras estruturas de visualização: 66% dos respondentes identificaram o gráfico de barras usando o termo correto para nomeá-lo, e 59% alcançaram o mesmo sucesso para o gráfico de linha (BÖRNER *et al.*, 2016). Essas conclusões não garantem acurácia na interpretação da informação por parte dos respondentes, mas sugerem alta capacidade de reconhecimento desses formatos.

Quanto ao mapa coroplético, apesar de familiar e encontrado nas ferramentas analisadas como solução dominante para a exibição de dados geoespaciais, trata-se de uma estrutura cujas desvantagens podem desencadear em informação enganosa. Mapas coropléticos exibem “atributos quantitativos codificados em cor sobre regiões delimitadas como marcadores de área” (MUNZNER, 2014, p. 181). Em todos os casos vistos nas ferramentas analisadas, os limites das regiões derivam de escolhas cartográficas e representam os limites administrativos de países, estados, ou municípios e condados. A codificação em cor presume um dado limite administrativo inteiramente com o mesmo valor, ocultando variações que podem existir ao longo da área. Além disso, limites de áreas pequenas podem não ser percebidos enquanto limites de áreas grandes se destacam.

Outras soluções de codificação de dados geospaciais que podem evitar essas desvantagens, como mapas de bolhas (*bubble map*), cartogramas (*cartogram*), *dot density maps* e *pin maps* foram identificadas nas ferramentas, porém em menor proporção.

O quinto padrão (tabela 2A (5)) destaca, por outro lado, exceções aos dois padrões anteriores. Estruturas de visualização menos corriqueiras, como linha do tempo com bolhas (*bubble timeline*), coordenadas paralelas (*parallel sets*), gráfico de área radial (*radial area chart*) e *strip plot* são alguns exemplos identificados no grupo de artigos de *scrollytelling*. É possível supor que esse formato de apresentação seja mais estimulante para o uso de estruturas menos usuais por duas razões. A primeira delas diz respeito ao meio de publicação: no jornalismo digital os periódicos disputam a atenção do leitor na web, o que corrobora com a hipótese de que visualizações mais atraentes e disruptivas sejam preferidas. A segunda diz respeito ao papel dos gráficos nesses artigos, que são comumente um suporte ao texto (SEYSER; ZEILLER, 2018), e não a fonte primária da informação. Nesse caso, pode haver uma menor preocupação com a acurácia da interpretação visual do que aquela encontrada no design de *dashboards*, com uma maior parte da segurança do entendimento da mensagem confiada ao texto.

O sexto padrão (tabela 2B (6)) revela a diferença na quantidade e diversidade de técnicas de interação também entre os dois formatos de apresentação. *Dashboards* oferecem uma gama de técnicas significativamente maior, não somente pela necessidade de fornecer ao usuário caminhos abertos de exploração, mas também pela abundância de variáveis, como já identificado. Conjuntos de dados grandes e complexos tornam o uso da interatividade crucial, pois é esse recurso que controla de que maneiras os usuários podem mudar o que eles visualizam (MUNZNER, 2014).

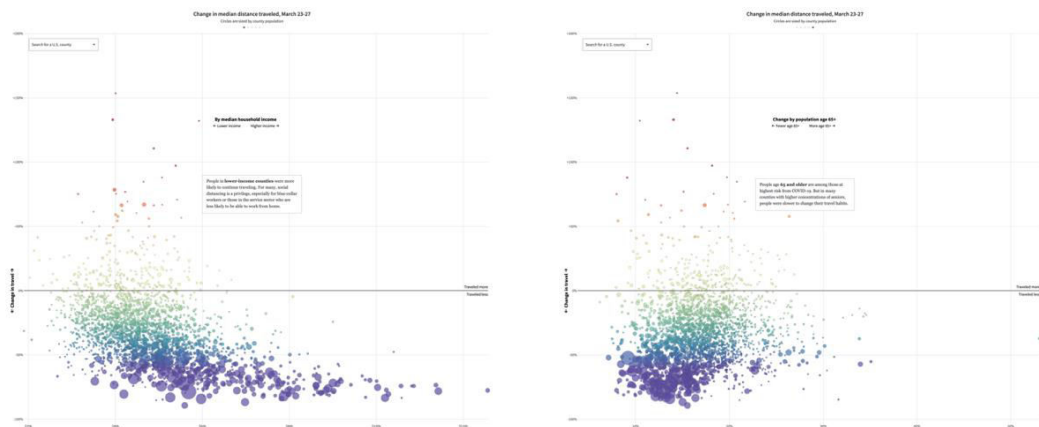
Por exemplo, uma ferramenta de visualização interativa pode dar suporte à investigação em vários níveis de detalhes, desde uma visão geral de alto nível, passando por vários níveis de sumarização, a até uma visão totalmente detalhada de uma pequena parte. Também pode apresentar diferentes formas de representar e resumir os dados de maneira que sustente a compreensão das conexões entre essas alternativas. (MUNZNER, 2014, p. 9).

Artigos de *scrollytelling*, por outro lado, apresentam os dados por uma estrutura narrativa, podendo ser linear ou elástica – essa última que permite caminhos ramificados (SEYSER; ZEILLER, 2018). Apesar de nenhuma das ferramentas desse grupo analisadas terem estrutura elástica, mesmo quando ela ocorre ainda assim é através de sequências pré-determinadas. Essa previsão do caminho que o usuário deverá seguir, impossível em formatos voltados para exploração aberta como os *dashboards*, reduz a necessidade de diferentes técnicas de interação a somente o escopo daquelas que guiarão o usuário no entendimento da história. Por exemplo, técnicas do tipo *details-on-demand*, como as operações *tool-tip* e *focus and*

detail, são fornecidas como mecanismo básico de elaboração da informação apresentada em visão geral no infográfico e contada no texto.

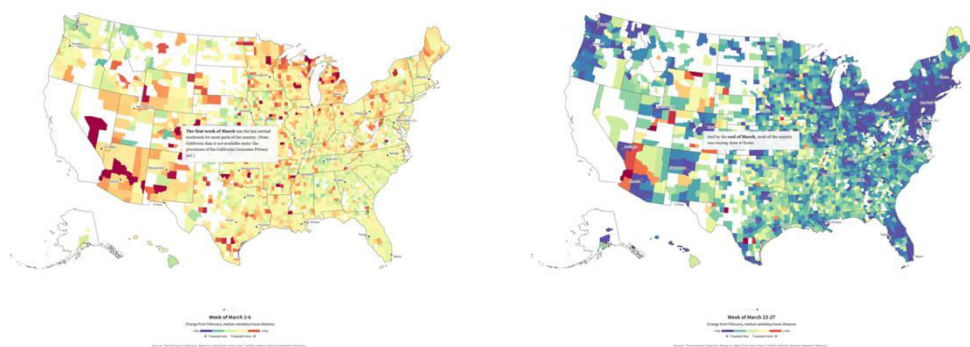
Uma característica específica da interatividade nos artigos de *scrollytelling* e que fica oculta na análise através da tabela 2 diz respeito ao operador das técnicas de interação. Comumente a interatividade é acessada pelo usuário através do clique ou do *hover* com o mouse, mas em artigos de *scrollytelling* algumas técnicas são engatilhadas pela rolagem (*scroll*). É o exemplo da ferramenta *The Social Distancing of America*, da Reuters. Nela, a operação de mudar atributos (*change attributes*), que permite reconfigurar a visualização trocando suas variáveis, é acionada em um gráfico de bolhas (*bubble chart*) não pelo clique do mouse, mas pelo *scroll* (Figuras 18 e 19). Uma operação de filtro – categorizada na ferramenta como outros controladores de consulta dinâmica (*other dynamic query controls*) – permite mudar a semana exibida em um mapa coroplético também através do *scroll* (Figuras 20 e 21). Nos dois casos, no entanto, a interação é restrita à sequência determinada pelo designer: você pode apenas ir e voltar numa sequência pré-definida de mudanças rolando a página para baixo e para cima, respectivamente.

Figuras 18 e 19: Capturas de tela exibindo a mudança de atributos plotados no gráfico engatilhada pelo *scroll* na ferramenta *The Social Distancing of America*.



Fonte: Reuters, 2020.

Figuras 20 e 21: Capturas de tela exibindo a troca de filtros de período aplicados no mapa engatilhada pelo *scroll* na ferramenta *The Social Distancing of America*.



Fonte: Reuters, 2020.

O sétimo padrão identificado (tabela 2B (7)) destaca a quase onipresença de técnicas de interação específicas nos *dashboards*, algumas delas avançadas. Na maioria dos gráficos encontrados nesse grupo é possível selecionar “itens de interesse nos dados para acompanhá-los” (YI et al., 2007) usando a operação de marcar (*marking*); trocar as variáveis plotadas com a operação de mudar atributos (*change attributes*); e ampliar e reduzir a escala de exibição através do zoom geométrico (*geometric zoom*).

As técnicas de filtragem também são amplamente disponibilizadas nos *dashboards* e “permitem que os usuários alterem o conjunto de dados apresentados baseando-se em condições específicas” (YI et al., 2007). As condições escolhidas pelo usuário refazem a consulta (*query*) nos dados em tempo real (*dynamic queries*) e normalmente são selecionadas através de controladores na interface. Para fins práticos de análise, os diversos tipos de componentes de interface de usuário identificados como controladores de consulta dinâmica – como *check boxes*, *radio buttons*, *toggle buttons*, *toggle switches*, *choice chips*, *multi-select chips*, *selection inputs*, *alphaslidars*, *rangeslidars* e campos de busca – foram sumarizados nos grupos *slider controls*, *selection controls* e *other dynamic query controls*.

As técnicas de conexão, como as operações *brushing* e *highlight by hovering*, podem ser consideradas avançadas e são de enorme conveniência no domínio a que se dedica essa análise, pois permitem justamente “destacar associações e relações entre dados” (YI et al., 2007). *Brushing* e *highlight by hovering* podem identificar a representação correspondente de um dado ou a representação de outro dado relacionado entre visualizações diferentes (YI et al., 2007), o que é bastante útil e rápido para um usuário que queira inferir **relações entre variáveis**.

O oitavo e último padrão (Tabela 2B (8)) identifica o uso de plataformas comerciais de *business intelligence* (BI) e de sistemas de informação geográfica (GIS) na construção de alguns *dashboards*, em oposição ao desenvolvimento de software sob demanda. Produtos disponíveis no mercado como ArcGIS, da ESRI; Google Data Studio, da Google; PowerBI, da Microsoft; e Tableau, da empresa de mesmo nome, fornecem ferramentas pré-configuradas que auxiliam desde o tratamento dos dados até a elaboração de visualizações interativas.

4 Limitações

Algumas potenciais limitações a seguir são reconhecidas na pesquisa. Em primeiro, a amostra de ferramentas obtida pelos critérios de seleção é reduzida em termos de diversidade linguística e regional. Quanto aos idiomas, só foram considerados o português e o inglês, por serem de domínio fluente do autor. Quanto às regiões abordadas, em 23 das ferramentas elas pertencem aos Estados Unidos. Das 10 restantes, 5 abordam o mundo, 3 o Brasil, 1 a Bélgica e 1 o Panamá.

Em segundo, a amostra também se resume a ferramentas disponíveis na web. Esse critério foi eleito por garantir possibilidade de acesso. Desconhece-se, portanto, que escolhas de design são empregadas em ferramentas somente internas, como por exemplo painéis de monitoramento usados em órgãos de pesquisa e de gestão de saúde.

Por fim, apesar do esforço criterioso em identificar elementos prescritivos de projeto, reconhece-se que essas descobertas não são generalizáveis para todos os contextos. Tratam-se de referências ricas para o design de ferramentas de visualização, especialmente daquelas que buscam exibir relações entre eventos, mas não avaliadas em critérios qualitativos, como efetividade, usabilidade e acionabilidade.

5 Considerações finais

Esse artigo analisou 33 ferramentas de visualização de dados disponíveis na web que permitem observar relações entre eventos de COVID-19 e variáveis preexistentes à doença. Os padrões e regularidades de design encontrados na amostra analisada respondem a perguntas úteis na concepção e projeto de ferramentas do tipo, como por exemplo: (1) que dados são normalmente disponibilizados para serem relacionados?; (2) que estruturas são comumente empregadas para codificar visualmente esses dados?; (3) que funções estatísticas as visualizações presentes permitem aos usuários?; (4) que técnicas de interação estão disponíveis para auxiliar o usuário na exploração e entendimento de relações entre os dados?; e (5) que linguagens de programação, empresas desenvolvedoras e/ou plataformas são escolhidas para a publicação dessas ferramentas na web?

Apenas as observações feitas na amostra analisada não são capazes de definir um conjunto de boas práticas para o design de ferramentas do tipo, uma vez que essa intenção foge do escopo da pesquisa. Entretanto, acredita-se na contribuição desse artigo com elementos prescritivos que possam auxiliar a comunidade do design da informação e da visualização de dados que visa produzir visualizações efetivas no entendimento de correlações por parte do público leitor.

6 Referências bibliográficas

- BÖRNER, K. et al. Investigating aspects of data visualization literacy using 20 information visualizations and 273 science museum visitors. **Information Visualization**, v. 15, n. 3, p. 198–213, 1 jul. 2016.
- CAIRO, A. **How Charts Lie: Getting Smarter about Visual Information**. [s.l.] W. W. Norton & Company, 2020.
- CIUCCARELLI, P.; KAHN, P. From scientific visualization to public engagement. Learning from a public archive of COVID-19 related visualizations. **Malofiej Special Covid-19**, n. 28, p. 7, 2021.
- FERDIO. **Data Viz Project | Collection of data visualizations to get inspired and finding the right type**. Disponível em: <<https://datavizproject.com/>>. Acesso em: 7 abr. 2022.
- FERREIRA, M. C. Spatial association between the incidence rate of COVID-19 and poverty in the São Paulo municipality, Brazil. **Geospatial Health**, v. 15, n. 2, 26 nov. 2020.
- FEW, S. **Information Dashboard Design: The Effective Visual Communication of Data**. 1st ed ed. Beijing ; Cambride [MA]: O'Reilly, 2006.
- GIANNELLA, J. R.; VELHO, L. Data visualization in the time of coronavirus. **Strategic Design Research Journal**, v. 14, n. 1, p. 275–288, 9 abr. 2021.
- HEER, J.; BOSTOCK, M.; OGIEVETSKY, V. A Tour Through the Visualization Zoo. **Communications of the ACM**, v. 53, n. 6, p. 59–67, jun. 2010.
- LI, S. L. et al. Higher risk of death from COVID-19 in low-income and non-White populations of São Paulo, Brazil. **BMJ Global Health**, v. 6, n. 4, p. e004959, 1 abr. 2021.
- MUNZNER, T. **Visualization Analysis and Design: Principles, Techniques, and Practice**. 1ª edição ed. Boca Raton: AK Peters, 2014.
- SEYSER, D.; ZEILLER, M. **Scrollytelling – An Analysis of Visual Storytelling in Online Journalism**.

2018 22nd International Conference Information Visualisation (IV). **Anais...** Em: 2018 22ND INTERNATIONAL CONFERENCE INFORMATION VISUALISATION (IV). jul. 2018.

SHNEIDERMAN, B. **The eyes have it: a task by data type taxonomy for information visualizations**. Proceedings 1996 IEEE Symposium on Visual Languages. **Anais...** Em: PROCEEDINGS 1996 IEEE SYMPOSIUM ON VISUAL LANGUAGES. set. 1996.

THEBAULT, R.; TRAN, A. B.; WILLIAMS, V. The coronavirus is infecting and killing black Americans at an alarmingly high rate. **Washington Post**, 7 abr. 2020.

UNSW SIDNEY. **Types of Data & the Scales of Measurement | UNSW Online**. Disponível em: <<https://studyonline.unsw.edu.au/blog/types-of-data>>. Acesso em: 15 abr. 2022.

YI, J. S. et al. Toward a Deeper Understanding of the Role of Interaction in Information Visualization. **IEEE Transactions on Visualization and Computer Graphics**, v. 13, n. 6, p. 1224–1231, nov. 2007.