

Visualização de dados para problemas urbanos

As águas de São Paulo
e a emergência climática

Nicholas Zbigniew Pretto Kaminski

Orientador: Prof. Dr. Leandro Manuel Reis Velloso
Coorientadora: Prof. Dr.^a Denise Helena Silva Duarte

Trabalho final de graduação em Arquitetura e Urbanismo
São Paulo, FAU USP, julho de 2022

Aviso

Este trabalho final de graduação é composto por dois elementos: um caderno sobre a emergência climática e as águas de São Paulo, baseado em infografia, e uma monografia (este documento), que descreve a metodologia, as referências e discorre sobre os infográficos.

Este relatório é indissociável do caderno que o acompanha. Sugere-se a leitura do caderno antes deste relatório.

As bases de dados públicas usadas neste trabalho podem ser acessadas por um repositório no GitHub: <<https://github.com/nicholaspretto/TFG>>.

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Catálogo na Publicação
Serviço Técnico de Biblioteca
Faculdade de Arquitetura e Urbanismo da Universidade de São Paulo

Kaminski, Nicholas Zbigniew Pretto
Visualização de Dados para Problemas Urbanos: as águas de São Paulo e a emergência climática / Nicholas Zbigniew Pretto Kaminski; orientador Leandro Manuel Reis Velloso. coorientador Denise Helena Silva Duarte - São Paulo, 2022.
2 v..

Trabalho Final de Graduação (Bacharelado em Arquitetura e Urbanismo) - Faculdade de Arquitetura e Urbanismo da Universidade de São Paulo.

1. Dataviz. 2. Mudança Climática. 3. Infografia. 4. São Paulo. I. Velloso, Leandro Manuel Reis, orient. II. Duarte, Denise Helena Silva, coorient. III. Título.

Elaborada eletronicamente através do formulário disponível em: <<http://www.fau.usp.br/fichacatalogica>>

Agradecimentos

Minha trajetória na graduação foi longa. Seria ingênuo supor que o trabalho final não viria acompanhado de uma lista de agradecimentos também longa.

Primeiramente, preciso agradecer à minha mãe, Landemir, e ao meu pai, Lucio, pelo amor, carinho e suporte que tive desde sempre. Não consigo mensurar o papel gigantesco que vocês tiveram na minha formação enquanto pessoa e enquanto profissional. Amo vocês.

Gostaria de agradecer a todos os meus amigos de Porto Alegre. A cidade em que eu não nasci, mas me fiz indivíduo e cidadão, também me presenteou com pessoas por quem sou extremamente grato por conhecer e manter contato. Ficam agradecimentos especiais à Nicole Freitas, minha grande companheira no curso da UFRGS e amiga pra todas as horas; à Morgana Goulart, que toda vez que conversamos me faz lembrar de como a Arquitetura e o Urbanismo são apaixonantes; e à Drika Garcia, com quem tenho uma conexão muito forte.

Os anos em São Paulo, especialmente os primeiros, não foram tranquilos. Justamente por isso, tenho que agradecer imensamente às pessoas que contribuíram para que eu me sentisse mais acolhido nessa cidade. Ficam agradecimentos especiais ao Eric Dick, com quem morei no primeiro ano e é um grande amigo; ao Samuel Bicego, uma das pessoas mais brilhantes que conheço e uma das amizades mais importantes que fiz em SP; e ao Daniel Alves, companheiro nas discussões sobre transporte público que se tornou um dos meus melhores amigos e que esteve comigo em momentos-chave, inclusive durante esse trabalho.

Na academia, não posso deixar de agradecer à Prof.^a Luísa Rocca, da UFRGS, por sempre acreditar em mim; à Prof.^a Luciana Royer, que me acolheu na sua pesquisa naquele assustador primeiro ano de FAU; e ao Giuseppe Filocomo, que foi importante na minha introdução ao mundo dos dados.

Meu primeiro estágio estudando na FAU foi no Metrô de São Paulo. Agradeço à equipe do CAC pela preocupação com meu bem-estar e aprendizado, e por ter me estimulado a trabalhar na Comunicação Visual, experiência que despertou meu interesse pelo design da informação. Um agradecimento especial à Juliana Tiemi, arquiteta incrível que me acompanhou e me ensinou muita coisa.

Minha entrada no mundo do dataviz foi pela cartografia. Agradeço às professoras de Planejamento Urbano da FAU pela introdução ao geoprocessamento e pelo estímulo a questionar “que história esse mapa quer contar?”, que tenho como mantra até hoje. Agradeço ao Leonardo Marques, não só por ter sido meu principal instrutor de QGIS, mas também por acreditar em mim e ter proporcionado meus primeiros trabalhos com cartografia. Sou grato, também, às pessoas que gostam dos mapas e gráficos que passei a publicar no Instagram e me estimularam a produzir mais e a ter orgulho do que faço.

A oportunidade de trabalhar no Nexa Jornal apareceu na segunda metade de 2021 e foi o que me permitiu ter contato diário com a produção de dataviz. Apesar de me sentir um peixe fora d'água no início, com dificuldades em uma variedade de frentes muito maior do que nos meus outros estágios, fui progressivamente me adaptando ao trabalho e às discussões do jornalismo. Sou grato pelo rico espaço de aprendizado proporcionado pela equipe de Gráficos, seja com escrita de materiais, seja com programação em R, seja pelas possibilidades de discussão sobre dataviz, seja na inserção no universo da produção das notícias. Sou grato também pelas amizades no caminho.

Agradeço ao Prof. Leandro Velloso por aceitar orientar esse trabalho mesmo sem me conhecer antes. Sou grato pelo direcionamento, pelos puxões de orelha quando a produção não avançava, pelas clareadas de ideias quando estava perdido e pela indicação de bibliografia, que me permitiu discutir e pensar sobre visualização de dados com maior embasamento teórico.

Agradeço também à Prof.^a Denise Duarte pela coorientação, termo que subestima sua contribuição para que o trabalho fosse possível. Sem as palavras de estímulo da professora, eu jamais teria me aprofundado tanto no tema das mudanças climáticas. Obrigado pelas indicações e correções sempre muito pertinentes e também por me fazer entender que minhas angústias do TFG não eram de outro mundo.

Não posso deixar de agradecer a todas as pessoas com quem pude trocar ideias e desabafar inseguranças durante o processo. Daniel Alves, Gabriel Maia, Gabriella Sales, Lucas Almeida, Luiz Salles e Morgana Goulart, muito obrigado por terem se disposto a gastar um tempo de vocês vendo meus infográficos, me ouvindo choramingar nos momentos de crise e me incentivando quando necessário. Agradeço também ao Rodolfo Almeida pela conversa muito elucidadora sobre infografia de mudanças climáticas e por ter me ajudado a entender o que eu estava buscando fazer: uma “história em gráficos”.

Agradeço à infografista e ilustradora Catarina Pignato e ao Prof. Eugenio Queiroga por terem aceitado compor essa banca. Sou grato pelo precioso tempo dispendido com a leitura deste trabalho e pelos seus comentários.

Finalmente, agradeço ao esforço monumental de formadores de políticas públicas, jornalistas, programadores e a sociedade civil organizada pelas iniciativas de disponibilização de dados públicos. Quando temos um primeiro contato na graduação com iniciativas como o Geosampa, não fazemos ideia de quão recentes e localizadas elas são. Fragilidades recentes em sistemas federais, marcadamente dados relativos à pandemia de Covid-19, também revelam que a disponibilidade de dados não é uma certeza eterna. Devemos buscar o fortalecimento da Lei de Acesso à Informação e tratar os dados como patrimônio público, que deve ser divulgado e preservado.

Resumo

O objetivo do trabalho é explorar o uso de dados públicos para contar histórias de relevância arquitetônica e urbanística, utilizando a visualização de dados, ou dataviz, como ferramenta. O produto final é constituído por duas partes: uma monografia, que inclui as definições de visualização de dados e jornalismo de dados, assim como discussões sobre cada um dos infográficos desenvolvidos; e um caderno, baseado em infografia, sobre a relação entre as águas da cidade de São Paulo e o impacto e resiliência da área urbanizada frente às mudanças climáticas. A história é baseada em dados, utilizando recursos visuais como gráficos e mapas. As teorias de visualização de dados, o jornalismo de dados e a infografia são usados como base teórica na estruturação da narrativa e nas estratégias gráficas.

Termos-chave: visualização de dados, infografia, mudanças climáticas, São Paulo, gráficos

Abstract

The objective of this work is to explore the use of public datasets to tell stories of architectural and urbanistic relevance, using data visualization, or dataviz, as a tool. The final product is composed by two parts: a thesis, including the definitions of data visualization and data-driven journalism, as well as discussions on each infographic; and a book, based in infographics, about the relation between the waters of the city of São Paulo and the impact and resilience of urban areas towards climate change. The story is based on data, using visual resources such as charts and maps. The theories of data visualization, data-driven journalism and infographics are used as a theoretical framework in the structuration of the narrative and the graphic strategies.

Keywords: data visualization, infographics, climate change, São Paulo, charts

Índice de figuras

Figura 1 — Mapa com a média móvel de mortes por Covid-19 no Brasil sendo mostrado no Jornal Nacional. Pág. 9

Figura 2 — Gráfico de linhas com casos de Covid-19 por região do mundo. Pág. 9

Figura 3 — Gráfico do Sexto Relatório do IPCC (AR6) com as regiões em que já se observa calor extremo. Pág. 10

Figura 4 — Mapa com mudanças na probabilidade de ondas de calor na América do Sul. Pág. 11

Figura 5 — Trecho de reportagem sobre o impacto das mudanças climáticas em Chicago. Pág. 12

Figura 6 — Mural de Catalhöyük citado no Milestones Project. Pág. 13

Figura 7 — Gráfico mostrando a frequência de uso, em literatura em língua inglesa, dos termos visualização de dados, infografia, design da informação e dataviz. Pág. 14

Figura 8 — O Censo é a pesquisa estatística mais importante do país. Pág. 15

Figura 9 — Ciclo de um projeto típico de ciência de dados. Pág. 16

Figura 10 — Uso do Illustrator no curso “Visualização de dados e design de informação”, da designer italiana Federica Fragapane. Pág. 17

Figura 11 — Versão impressa do South China Morning Post com infográfico sobre mudanças urbanas pós-pandemia. Pág. 18

Figura 12 — Parte do material interativo sobre o incêndio na Catedral de Notre Dame de Paris em 2019. Pág. 19

Figura 13 — Gráfico do jornal Nexo que não seria chamado “infográfico” pela definição de Teixeira (2007), já que conta, sozinho, uma história. Pág. 21

Figura 14 — Tabela proposta para mensurar reportagens numa escala de “totalmente jornalismo com dados” até “totalmente jornalismo de dados”. Pág. 25

Figura 15 — Trecho da reportagem sobre Chicago e a emergência climática. Pág. 26

Figura 16 — Parte de planilha com os datasets encontrados para Buenos Aires, obtidos por raspagem do site Buenos Aires Data. Pág. 27

Figura 17 — Tabela com as tipografias de jornais selecionados. Pág. 30

Figura 18 — Diagrama tipo Sankey mostrando o destino das exportações de petróleo da Rússia. Pág. 32

Figura 19 — Diagrama de fluxo proposto por Charles Minard. Pág. 33

Figura 20 — Mapas com os fluxos de gás natural russo para a Europa. Pág. 34

Figura 21 — Uma tarde de domingo na ilha de Grande Jatte, de Georges Seurat. Pág. 35

Figura 22 — Mapa proposto por Burney. Pág. 36

Figura 23 — Mapa de Fortaleza utilizado como exemplo no portal Acesso a Oportunidades. Pág. 36

Figura 24 — Zoom na região do município de São Paulo. Pág. 37

Figura 25 — Mapa dos rios da Turquia. Pág. 38

Figura 26 — Parte do mosaico de folhas do mapa Sara Brasil. Pág. 39

Figura 27 — Mapa de São Paulo em 1842. Pág. 40

Figura 28 — Primeira versão do mapa. Pág. 41

Figura 29 — Mapa da região dos Grandes Lagos, destacando as massas d’água. Pág. 42

Figura 30 — Proposta inicial para mapa de expansão urbana. Pág. 43

Figura 31 — Segunda versão do infográfico de expansão urbana. Pág. 44

Figura 32 — Gráfico “cru” incluindo a variação entre 1997 e 2002 dentro do dataset. Pág. 44

Figura 33 — Tableau Poléométrique, de Fourcroy. Pág. 45

Figura 34 — Cartograma retangular com a população dos EUA em 1930. Pág. 46

Figura 35 — Treemap de exportações do Brasil em 2020, em dólares. Pág. 47

Figura 36 — Entorno da Ponte Estaiada. Pág. 48

Figura 37 — Mapa de John Snow. Pág. 49

Figura 38 — Mapa Racial do Brasil numa escala afastada. Pág. 50

Figura 39 — Gráfico waffle. Pág. 51

Figura 40 — Composição da Assembleia Nacional Francesa por partido. Pág. 51

Figura 41 — Gráfico de Buache e L'Isle. Pág. 52

Figura 42 — Gráfico de barras de Playfair. Pág. 53

Figura 43 — Infográfico de Scarr. Pág. 54.

Figura 44 — Infográfico de Arranz e Huang. Pág. 55

Figura 45 — Gráfico de Playfair mostrando a balança comercial inglesa. Pág. 56

Figura 46 — Gráfico de emissões de CO2 por combustíveis fósseis. Pág. 57

Figura 47 — Infográfico dividido em cores sólidas na horizontal. Pág. 57

Figura 48 — Cachecol original de Sheldon. Pág. 58

Figura 49 — Cobertor original de Highwood. Pág. 58

Figura 50 — Primeiras listras climáticas de Ed Hawkins. Pág. 59

Figura 51 — Listras climáticas para o Brasil. Pág. 59

Figura 52 — Modelo na London Fashion Week usando roupa com listras climáticas. Pág. 60

Figura 53 — Trecho do gráfico original do Listening History. Pág. 61

Figura 54 — Trecho do gráfico do New York Times mostrando a arrecadação de filmes nos EUA. Pág. 62

Figura 55 — Trecho de streamgraph mostrando o número de apartamentos novos por região de São Paulo, por ano. Pág. 63

Figura 56 — Versão inicial do mapa de espaços livres. Pág. 65

Figura 57 — Página da biografia gráfica de São Paulo com os espaços livres. Pág. 66

Figura 58 — Mapa de rios e córregos proposto pelo G1. Pág. 67

Figura 59 — Gráfico original sobre adaptação urbana. Pág. 69

Sumário

1. Introdução	9	5. Considerações finais	70
2. Conceitos	13	5.1. Lições aprendidas	70
2.1. Visualização de dados	13	5.2. Conclusão	72
2.2. Infografia	18	6. Referências bibliográficas	73
2.3. Jornalismo de dados	22		
3. Métodos	26		
3.1. TFG 1 (agosto a dezembro de 2021)	26		
3.2. TFG 2 (março a julho de 2022)	28		
3.3. Guia de estilo	29		
4. Fichas dos infográficos	31		
Infográfico 1: Benevolo	32		
Infográfico 2: Américas	35		
Infográfico 3: Bacia do Prata	37		
Infográfico 4: Retificações	39		
Infográfico 5: Reservatórios	41		
Infográfico 6: Mancha urbana	43		
Infográfico 7: Desastres	45		
Infográfico 8: Aglomerados subnormais	48		
Infográficos 9 e 10: Chuvas	52		
Infográfico 11: Cantareira	56		
Infográfico 12: Warming Stripes	58		
Infográfico 13: Emissões	61		
Infográficos 14 e 15: Projeta	64		
Infográfico 16: Espaços livres	65		
Infográfico 17: Bacias de São Paulo	67		
Infográfico 18: Resiliência urbana	68		

1. Introdução

No começo de 2021, comecei a estudar de maneira mais aprofundada técnicas de cartografia e, principalmente, o uso do software QGIS. A partir dessas explorações e do contato com outras pessoas que trabalham com cartografia, conheci uma grande área chamada visualização de dados, ou “dataviz”.

Dataviz consiste em transformar dados em “visualizações”, ou seja, trazer uma narrativa visual a informações, geralmente tabelas. Os formatos mais habituais de dataviz são gráficos (como de barras, de linhas, de pizza) e mapas (que podem ser entendidos como “gráficos georreferenciados”).

O papel da visualização de dados se mostrou essencial durante a pandemia de Covid-19. Os gráficos buscam traduzir as informações disponibilizadas por governos e órgãos de saúde para um formato de mais fácil assimilação. Esse trabalho se mostrou de extrema importância para que a população compreendesse a gravidade da situação o mais rapidamente possível, transformando números frios em imagens impactantes (figuras 1 e 2).

Figura 1: Mapa com a média móvel de mortes por Covid-19 no Brasil sendo mostrado no Jornal Nacional. Fonte: G1. Disponível em <<https://g1.globo.com/jornal-nacional/noticia/2020/11/18/media-movel-de-mortes-e-casos-por-covid-segue-em-fortetendencia-de-alta-no-brasil.ghml>>. Acesso em 14 de novembro de 2021.

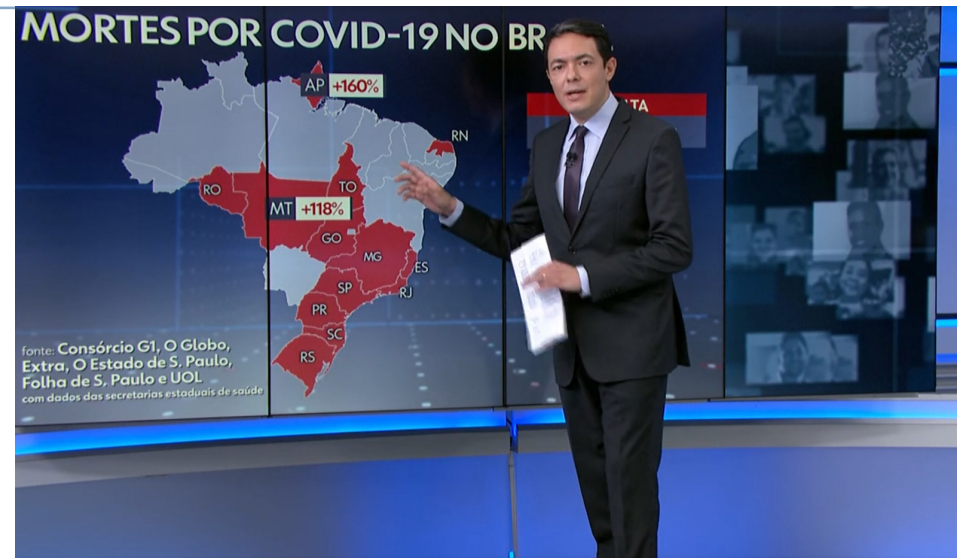
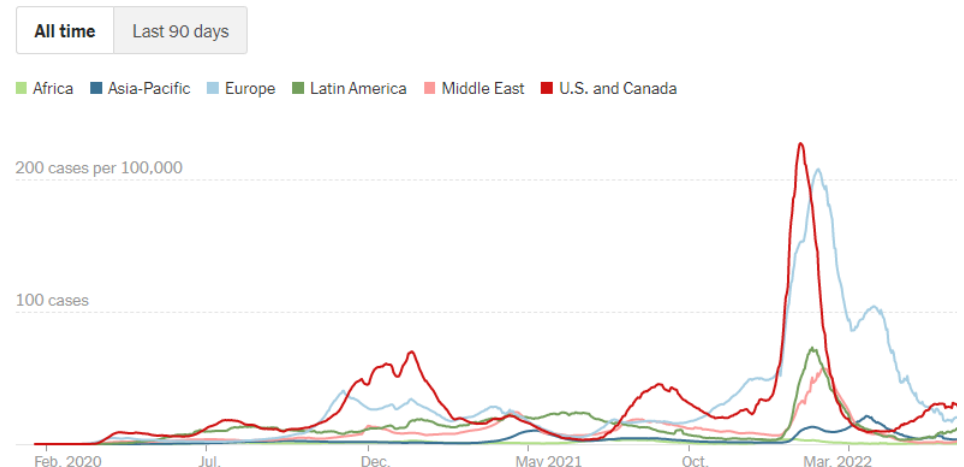


Figura 2: Gráfico de linhas com casos de Covid-19 por região do mundo. Fonte: The New York Times. Disponível em <<https://www.nytimes.com/interactive/2021/world/covid-cases.html>>. Acesso em 11 de julho de 2022.

Cases by region

This chart shows how cases per capita have changed in different parts of the world.



A visualização de dados é uma ferramenta de relevância para várias áreas de atuação mas, como visto nos exemplos anteriormente citados, é perceptível que se destaca no campo do jornalismo. O termo “jornalismo de dados” se refere à prática jornalística que busca noticiar ou reportar com base em dados, o que pode acontecer diretamente no texto corrido ou através de visualizações. É comum que essas últimas sejam utilizadas atreladas a matérias que se valham de mais imagens e diagramas para explicar determinado assunto ou fenômeno, ou seja, são parte importante de matérias de infografia.

Em julho de 2021, entrei para a seção de Gráficos do Nexo (<https://www.nexojournal.com.br/grafico/>), um jornal online que se destaca pelo uso de infografia para explicar assuntos. A área, no momento, é dividida em duas partes: a equipe de Dados e a de Infografia. A primeira é responsável por encontrar bases de dados, trabalhar em cima delas, produzir gráficos e mapas iniciais (utilizando principalmente a linguagem R e o software QGIS) e escrever histórias a partir do material. Já a segunda é responsável por transformar visualmente esse material, fazer o tratamento visual de todos os gráficos e mapas, produzir ilustrações adicionais e subir o material na plataforma de publicação. Atualmente, sou estagiário na equipe de Dados.

O jornalismo de dados brasileiro apresenta poucos profissionais formados em Arquitetura e Urbanismo, mas é curioso perceber como as discussões e métodos existentes em visualização de dados estão diretamente relacionados a debates comuns à nossa graduação.

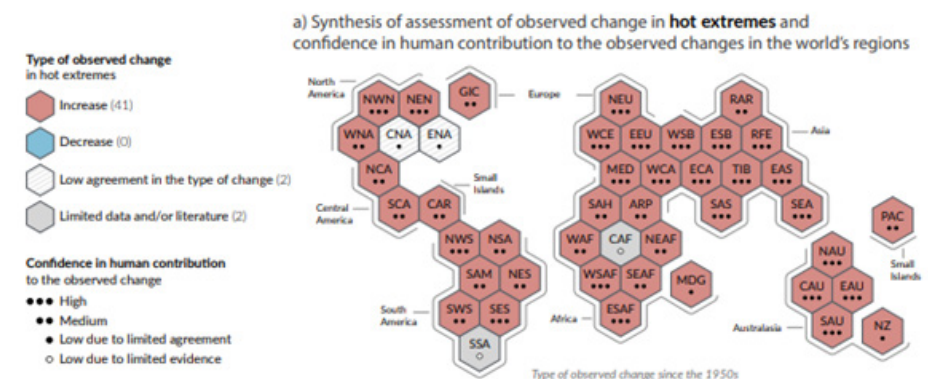
Uma planta pode ser compreendida como uma espécie de dataviz: transforma informações espaciais (nas quais atributos numéricos das dimensões são importantes) em um material visual de fácil compreensão. Os debates sobre representação das informações de modo belo e legível também são extremamente presentes em Arquitetura e Urbanismo, sempre atreladas a limites impostos pela informação a ser apresentada: a beleza importa, mas explicar o material sem ambiguidades é o mais importante. Os mapas produzidos nas disciplinas de Planejamento Urbano estimulam diálogos que, de maneira inconsciente, já atuam como pontapé inicial em discussões de como traduzir dados em informação visual. A visualização de dados em contexto jornalístico, portanto, é um campo profissional que pode ser desbravado também por arquitetos e urbanistas.

A visualização de dados contribui para que temas difíceis sejam explicados à população de maneira direta e de mais simples compreensão. As alterações climáticas que vêm sendo provocadas pela emissão de gases estufa correspondem a um conjunto de pautas que não apenas são urgentes ao debate público como exigem a compreensão de bases de dados complexas.

O trabalho de dataviz e de infografia, portanto, também é de imensa relevância para dar suporte ao trabalho de cientistas e para reforçar o tom de emergência que a crise exige. Exemplo desse uso são os gráficos e mapas presentes nos relatórios do IPCC (Painel Intergovernamental sobre Mudanças Climáticas das Nações Unidas) (figura 3).

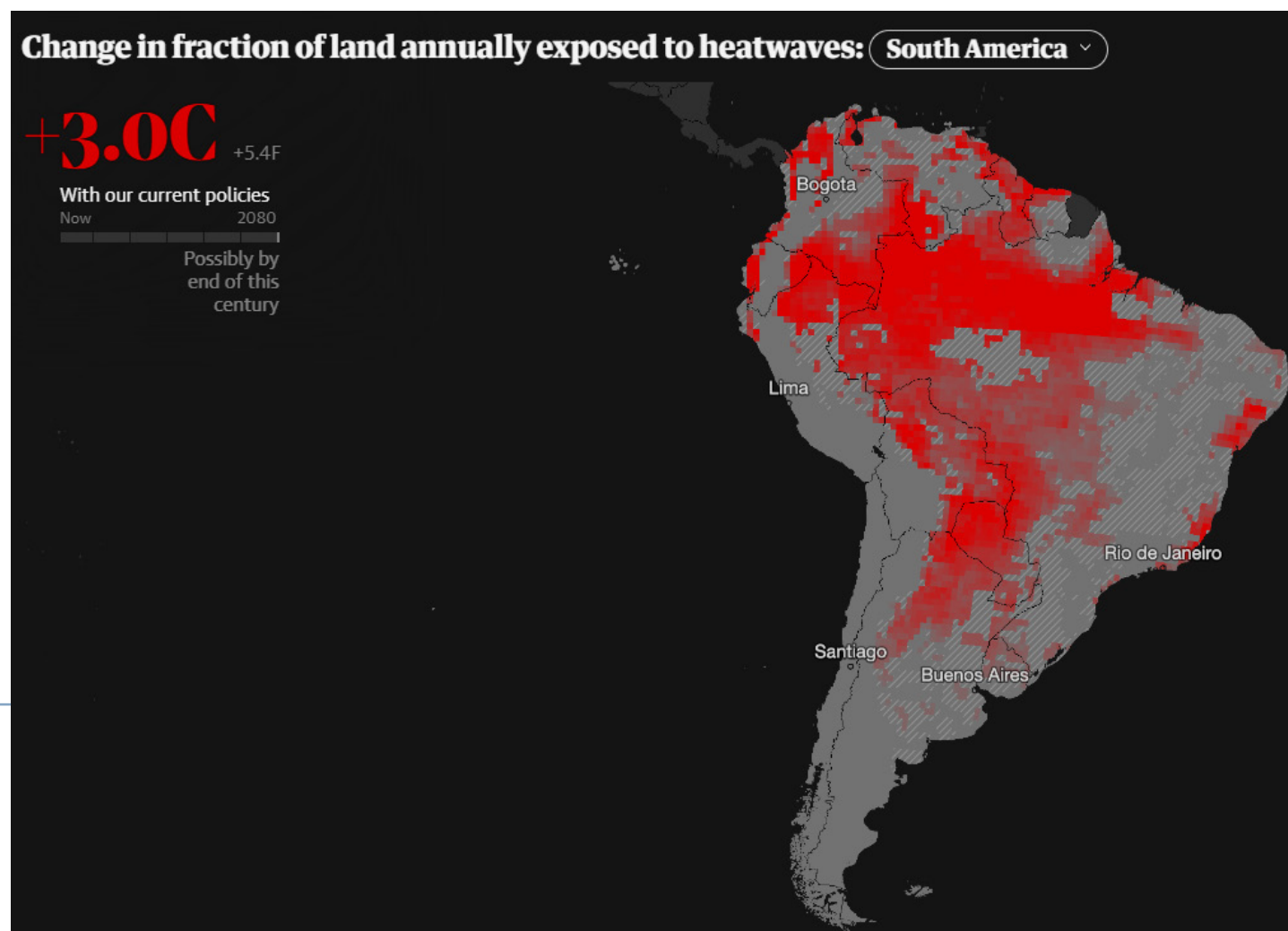
Figura 3: Gráfico do Sexto Relatório do IPCC (AR6) com as regiões em que já se observa calor extremo.
Fonte: IPCC. AR6 WGI Report. Cambridge: Cambridge University Press, 2021. Disponível em: <https://www.ipcc.ch/report/ar6/wg2/downloads/report/IPCC_AR6_WGII_FinalDraft_FullReport.pdf>. Acesso em 14 de novembro de 2022.

Climate change is already affecting every inhabited region across the globe with human influence contributing to many observed changes in weather and climate extremes



Apesar de importantes e com visualizações cada vez mais elaboradas (inclusive um atlas interativo no Sexto Relatório), os gráficos e mapas do IPCC não são sempre de fácil leitura para o cidadão comum. O tamanho dos relatórios, a complexidade do tema e a limitação do idioma (os relatórios completos são apenas apresentados em inglês) acabam estimulando que veículos de comunicação produzam seus próprios materiais derivados das informações dos trabalhos do IPCC (ver figura 4).

Figura 4: Mapa com mudanças na probabilidade de ondas de calor na América do Sul. Fonte: The Guardian. Disponível em <<https://www.theguardian.com/environment/ng-interactive/2021/oct/14/climate-change-happening-now-stats-graphs-maps-cop26>>. Acesso em 19 de novembro de 2021.



Uma lacuna em visualização de dados sobre mudanças climáticas, porém, se encontra no tratamento dado ao impacto da crise sobre as cidades de maneira específica. Tanto nos relatórios oficiais do IPCC quanto numa grande gama de reportagens sobre o tema, a abordagem costuma ser centrada nos impactos sobre o planeta como um todo, sobre grandes áreas dos continentes ou sobre determinados biomas (por exemplo, a Antártica, a Amazônia, o Cerrado etc.).

No Sexto Relatório do IPCC (AR6), como parte do Grupo de Trabalho II, há um capítulo específico que trata de cidades (capítulo 6, *Cities, settlements and key infrastructure*), mas ainda de forma muito ampla. Reportagens que abordam cidades específicas costumam se basear, principalmente, em entrevistas, pesquisa histórica e vídeos (figura 5); esses recursos são importantes, mas revelam a lacuna no uso de dataviz e sua potencialidade.



The morning of Jan. 11, 2020, on the lakeshore. Deborah Harrington

Figura 5: Trecho de reportagem sobre o impacto das mudanças climáticas em Chicago. Fonte: The New York Times. Disponível em: <https://www.nytimes.com/interactive/2021/07/07/climate/chicago-river-lake-michigan.html>. Acesso em 19 de novembro de 2021.

A abordagem municipal sobre alterações climáticas é de suma importância para a conscientização da população e também para o desenvolvimento de políticas públicas. Com a maior parte da população mundial vivendo em cidades, pensar em medidas adaptativas é essencial.

A partir de referências, pesquisas sobre a emergência climática e seu impacto nas cidades, bem como a análise dos bancos de dados disponíveis, o intuito cabal deste trabalho é traduzir essa abundância de informações através de uma narrativa com infografia. Esse processo envolve a elaboração de visualizações, o desenvolvimento de uma narrativa, o tratamento visual do conjunto e a disponibilização do trabalho por algum meio de fácil acesso.

O objetivo final deste trabalho, portanto, é revisitar o debate do impacto da emergência climática em São Paulo através de um material ilustrado, com embasamento sólido em fontes teóricas e uma narrativa geral que unifique distintas fontes de dados.

2. Conceitos

2.1. Visualização de dados

A SEDG (Society for Experiential Graphic Design) define design da informação da seguinte forma, em tradução livre:

Design da informação é a prática de apresentar a informação de modo a torná-la mais acessível e de fácil compreensão para os usuários. O design da informação é voltado a audiências específicas para atender objetivos definidos. Em suas formas mais sofisticadas, ajuda os usuários a entender dados complexos através da organização e simplificação da informação. (SEDG, n.d.)

A definição de design da informação da SEDG adota um juízo de valor ao falar sobre “formas mais sofisticadas” que abordem o uso de dados para explicar informação. É a partir dessa noção que entramos no universo dos gráficos e, como consequência, da “visualização de dados” (também comumente chamada pelo seu apelido em inglês, *dataviz*). O dicionário Cambridge dá a seguinte definição para “data visualization”, aqui em tradução própria:

Ato de representar a informação como uma imagem, diagrama ou gráfico, ou imagem que representa a informação dessa maneira. (CAMBRIDGE DICTIONARY, n.d.)

A representação pictórica de dados tem precedentes já na pré-história (como em pinturas rupestres que utilizam formas em sequência para catalogar uma quantidade de objetos ou animais). O Milestones Project, que busca catalogar marcos importantes da visualização de dados, começa sua listagem com um mapa da cidade de Catalhöyük de 8000 anos atrás (FRIENDLY, 2021). Apesar disso, a sistematização do uso de gráficos é um fenômeno recente (tradução própria):

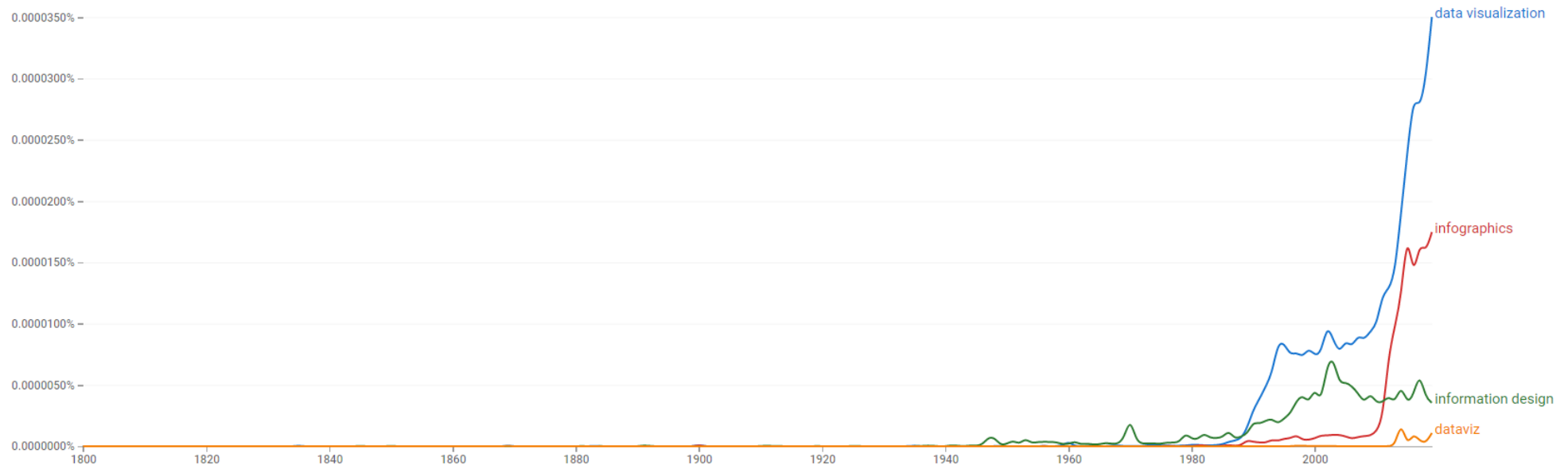
O uso de formas abstratas e não representativas para mostrar números são uma invenção surpreendentemente recente, talvez pela diversidade de habilidades exigida — a visual-artística, a empírica-estatística e a matemática. Foi só em 1750-1800 que gráficos estatísticos — largura e área para mostrar quantidades, séries temporais, gráficos de dispersão e visualizações de mais de uma variável — foram inventados, muito depois dos triunfos de engenhosidade matemática como logaritmos, coordenadas cartesianas, cálculo e as bases da teoria probabilística. O memorável William Playfair (1759-1823) desenvolveu ou melhorou quase todos os designs de gráficos fundamentais, buscando substituir tabelas numéricas convencionais com uma representação visual sistemática de sua “aritmética linear”. (TUFT, 2007, p.10)

Figura 6: Mural de Catalhöyük citado no Milestones Project. Fonte: Mapping as a Process. Disponível em: <https://www.mappingasprocess.net/blog/2017/12/14/this-is-not-a-map>. Acesso em 4 de julho de 2022.



O Google Books possui uma ferramenta chamada Ngram Viewer, que permite verificar a frequência de termos em livros de diversos idiomas (até o momento, sem versão em português). O gráfico de linhas a seguir (figura 7) aborda alguns dos termos discutidos neste trabalho. O eixo vertical representa a porcentagem de vezes em que o termo é encontrado em publicações em inglês daquele ano.

Figura 7: Gráfico mostrando a frequência de uso, em literatura em língua inglesa, dos termos visualização de dados, infografia, design da informação e dataviz. Fonte: Google Ngram Viewer. Disponível em <<https://books.google.com/ngrams>>. Acesso em 28 de novembro de 2021.



É evidente que, desde os anos 1990, a visualização de dados se tornou conceito dominante entre as terminologias da área. É difícil estabelecer relações de causalidade, mas o início da popularidade do termo na década de 1980 está provavelmente relacionado ao lançamento de *The Visual Display of Quantitative Information*, de Edward Tufte, em 1983.

Os termos “data visualization” e “information design” têm crescimentos similares entre 1997 e 2003, mas a partir desse momento, “information design” se mantém estável, enquanto “data visualization” cresce. Tanto “data visualization” quanto “infographics” têm uma explosão no uso a partir da década de 2010.

A explosão recente do interesse pela visualização de dados é diretamente influenciada por dois movimentos. Um deles é o aumento da quantidade de informações disponíveis que precisam ser repassadas com uma leitura fácil (principalmente ao lidar com quantidades enormes de dados, o chamado “big data”). O segundo é a velocidade de consumo da informação, potencializada pela internet e pelas redes sociais, que acaba privilegiando imagens que se destaquem. Cairo (2012) argumenta que nosso cérebro transforma toda a informação que recebemos em “representações administráveis”, ou símbolos; buscar a representação de dados de maneira visual, portanto, facilita a memória da mesma (CAIRO, 2012).

As diversas literaturas trazem argumentos favoráveis à tradução dos dados por meio de sua visualização. Os principais argumentos costumam circundar a ideia de que processamos imagens de maneira muito mais rápida do que tabelas. Em um episódio do podcast 99% Invisible, o economista Tim Harford cita um estudo do Tufts Visual Analytics Lab que diz que uma pessoa cria uma impressão sobre um gráfico em meio segundo (99% INVISIBLE, 2021). Esse tempo não é suficiente para formular uma ideia aprofundada acerca dos dados ali exibidos, mas é o suficiente para que o leitor tenha uma primeira opinião sobre ele (“bonito”, “feio”, “simples”, “complexo”), e também é determinante para que o leitor decida ou não se aprofundar na análise.

Ware (1999) cita cinco argumentos sobre as vantagens da visualização: permite compreender grandes quantidades de dados em pouco tempo; permite a percepção de propriedades novas por meio da observação de padrões; auxilia no controle de qualidade dos dados, pois a representação gráfica faz saltar aos olhos erros e valores muito discrepantes; facilita o entendimento tanto de características de grande quanto de pequena escala dos dados; e facilita a formulação de hipóteses (WARE, 1999).

Diante da relevância da visualização de dados, é de se esperar que seja uma habilidade almejada pelo mer-

cado de trabalho. Um ponto sensível a ser considerado, apesar disso, é a quantidade de competências exigidas envolvendo esforços cognitivos distintos. Um profissional com formação em design provavelmente não passou por muitas disciplinas de estatística durante a graduação; como contraponto, é comum vermos profissionais de diversas áreas que se valem da estatística apresentando gráficos totalmente ininteligíveis - ou, simplesmente, muito feios. No seu capítulo para o livro *Beautiful Visualization*, Noah Iliinsky argumenta que, para uma visualização ser esteticamente aprazível, ela deve ser inovadora, informativa e eficiente (STEELE e ILIINSKY, 2010).

Encontrar bases de dados e entender a metodologia de obtenção dos dados é, por si só, uma tarefa complexa: os bancos de dados são dispersos, metodologias mudam ao longo dos anos, e algumas informações simplesmente não existem. Não há como produzir um gráfico sobre a preferência dos brasileiros por certos tipos de flores se nunca foi feita uma pesquisa com uma amostra representativa da população questionando exatamente isso. A tentativa de criar correlações entre os resultados de uma pesquisa para que os dados corroborem com uma narrativa desejada é uma má prática na área.

É, também, importante refletir sobre a neutralidade da informação. Como dito anteriormente, as pessoas têm uma impressão inicial sobre um

gráfico muito rapidamente, tornando a responsabilidade ética na sua elaboração particularmente sensível. Assim, a utilização de dados confiáveis que busquem refletir a realidade é importantíssima: a neutralidade na coleta e na disponibilização de dados torna essencial o papel de institutos de estatística politicamente autônomos, no caso de informações públicas como o Censo (figura 8), e estimula a contratação de empresas independentes de pesquisa, no caso de corporações.

Figura 8: O Censo é a pesquisa estatística mais importante do país. Fonte: G1. Disponível em: <https://g1.globo.com/mg/zona-da-mata/concursos-e-emprego/noticia/2022/07/11/ibge-convoca-candidatos-aprovados-na-primeira-etapa-da-selecao-para-o-censo-2022-na-zona-da-mata-e-vertentes.ghml>. Acesso em 11 de julho de 2022.



Os dados obtidos geralmente exigem alguma espécie de tratamento. Dados que já se encontram “prontos” para serem transformados em gráficos, geralmente, foram tratados por alguém ou por alguma equipe (mesmo que seja a equipe do próprio instituto que divulga o dado). Assim, é interessante compreender as etapas desse tratamento. Para isso, se utilizará o chamado “ciclo da ciência de dados”, citado por Wickman e Grolemond (2017), como referência (figura 9).

A arrumação dos dados, ou “faxina”, consiste em armazená-los de maneira consistente e organizada. Isso significa, por exemplo, fazer com que cada coluna da tabela corresponda a uma variável: se existe uma coluna em que as cidades estão disponíveis como uma variável “Cidade, UF” e pretende-se utilizar somente “unidade da federação” como variável, será necessário dividir a coluna em duas (uma para o

nome da cidade, outra para a unidade da federação). Outra questão comum a ser resolvida é a presença de “x” ou “...” para representar valores ausentes em planilhas numéricas, o que acaba impossibilitando cálculo com as variáveis pela existência de texto não-numérico. Esse tratamento é importante para a utilização de bases de dados, e também costuma ocupar uma parcela expressiva de horas de trabalho.

A transformação dos dados consiste em trabalhar com eles de modo a prepará-los para a visualização. Um exemplo cotidiano de transformação de dados é o conceito de “tabela dinâmica” do Excel: a partir de uma base de dados maior, o programa permite agrupar variáveis e fazer filtros. Essas operações são especialmente importantes para a compreensão de grandes bases de dados: uma planilha gigantesca contendo a idade de milhares de indivíduos, por exemplo, é

humanamente impossível de ser compreendida no seu conjunto; a transformação em uma planilha que indica quantas pessoas de cada idade estão naquela amostra, porém, já é um indicador bem mais amigável. O cálculo de algumas métricas, como a média e a mediana de variáveis numéricas, também entram nessa etapa e nos auxiliam a ter uma dimensão melhor das informações disponíveis.

A modelagem dos dados é fundamental para a área de pesquisa e para o trabalho com ciência de dados em empresas. Um modelo econométrico, por exemplo, pretende conseguir elucidar relações entre variáveis e representar a realidade, buscando também explicar tendências futuras. No caso de estudos que exijam modelagem dos dados, a visualização também é muito importante, mas não necessariamente essa modelagem é essencial para a visualização. Quando a intenção é simplesmente elucidar os dados existentes, e não fazer inferências a partir de conjuntos de variáveis, essa fase não acontece.

A visualização dos dados é compreendida, dentro dessa proposta de ciclo de ciência de dados, como uma etapa desse processo circular de exploração, que engloba a transformação e a modelagem. Isso acontece pelas razões citadas anteriormente descritas por Ware (1999): a visualização possibilita ter insights e identificar erros nos estágios anteriores que não seriam perceptíveis de outra maneira. É comum que se faça diversos testes com formatos de gráficos diferentes até que se encontre a alternativa que melhor representa a narrativa que se deseja contar.

É interessante que, pelo ciclo descrito por Wickham e Grolemond (2007), o produto final não é chamado “visualização”, mas “comunicação”. Isso se deve ao fluxo de trabalho de um cientista de dados e pela própria área de atuação dos profissionais. No contexto empresarial, dificilmente a visualização dos dados vai ser compreendida como etapa final do processo: muito provavelmente, os gráficos farão parte de uma apresentação ou relatório. Alguns processos nas linguagens de programação permitem, inclusive, automatizar a atualização de documentos à medida que os dados são atualizados (ver o capítulo 27 de Wickham e Grolemond, 2007).

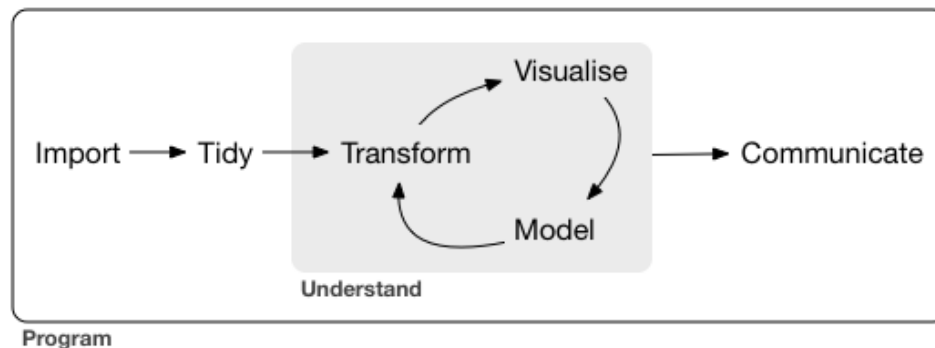


Figura 9: Ciclo de um projeto típico de ciência de dados. Fonte: Wickman e Grolemond (2017). Disponível em <<https://r4ds.had.co.nz/index.html>>. Acesso em 28 de novembro de 2021.

Já tendo um gráfico pronto, mas recém extraído de alguma plataforma sem tratamento estético mais refinado, é possível avançar para uma etapa de finalização. As linguagens de programação (como R e Python), plataformas de Business Intelligence (como Power BI e Tableau) e outros softwares (como Excel e QGIS) permitem a personalização de dados por meio da escolha de formas, cores, legendas, entre outros aspectos. Ainda assim, locais com uma equipe maior costumam utilizar softwares de edição de imagem e de vetores, como Illustrator e Photoshop, para fazer a finalização (figura 10).

Softwares de edição de imagem são interessantes para a visualização de dados, pois extrapolam a obviedade dos gráficos tradicionais. Cada uma das plataformas voltadas especificamente para visualização de dados apresenta limitações dadas pelas possibilidades fornecidas pela própria plataforma. Programas como o Illustrator entendem o gráfico simplesmente como uma imagem: se for necessário fazer um processo intermediário entre a importação do gráfico “cru” e sua entrega final, não haverá incompatibilidade, mesmo que a informação precise ser distorcida para depois ser reajustada.

É possível perceber, portanto, que o processo de visualização de dados não só envolve uma enorme gama de processos, mas também habilidades diversas e que não são obviamente complementares fora desse contexto. O melhor designer gráfico do mundo não necessariamente precisa saber programar; o melhor programador do mundo pode perfeitamente nunca ter aberto o Illustrator. Um profissional que pretenda participar de todas as etapas do dataviz, porém, precisa ter noções de todas as etapas anteriormente descritas. Por isso, é comum que empresas tenham equipes de profissionais de áreas distintas, às vezes com pouco contato entre si. O desafio envolvendo essa complexidade é trazido por Knaflitz (2019) no seu já clássico livro sobre dataviz em contexto corporativo:

Essa diversidade de habilidades também reflete a diversidade de possibilidades trazidas pelo uso da visualização de dados. Todas as áreas que precisam repassar informações a partir de bases de dados precisam se valer do tema para comunicar resultados. Este trabalho pretende focar na abordagem jornalística, por meio da infografia.

Figura 10: Uso do Illustrator no curso “Visualização de dados e design de informação”, da designer italiana Federica Fragapane. Fonte: Domestika. Disponível em: <https://www.domestika.org/pt/courses/2495-visualizacao-de-dados-e-design-de-informacao-crie-um-modelo-visual>. Acesso em 28 de novembro de 2021.



É difícil encontrar pessoas habilitadas na visualização de dados. Parte do desafio é que a visualização de dados é uma singular etapa do processo analítico. As pessoas contratadas para funções analíticas normalmente têm conhecimentos quantitativos, adequados para as outras etapas (encontrar os dados, reuni-los, analisá-los, construir modelos), mas não necessariamente algum treinamento formal em design para ajudá-los na comunicação da análise - o qual, a propósito, normalmente, é a única parte do processo que seu público vê. (KNAFLITZ, 2019, p.7)

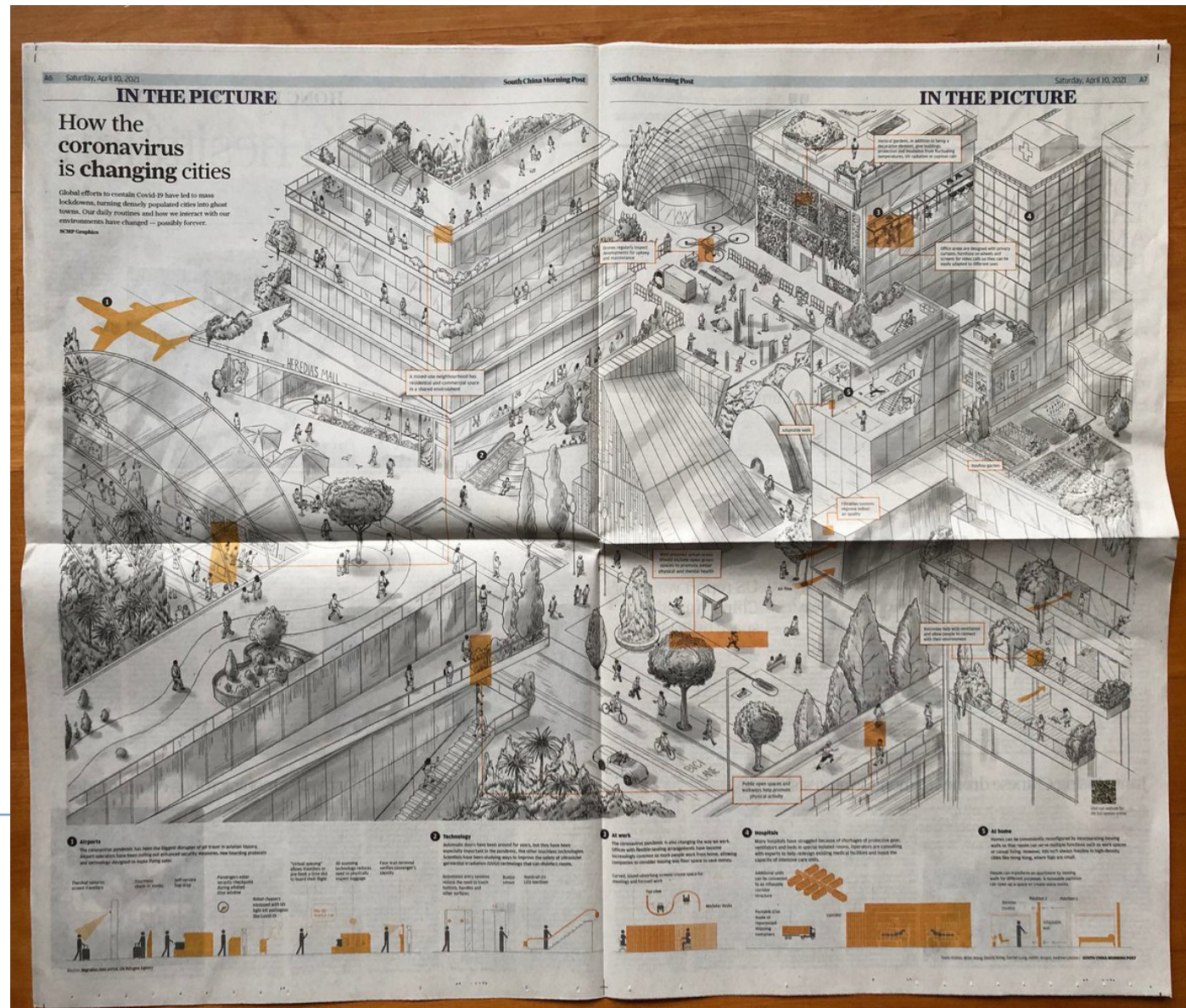
2.2. Infografia

Infografia é um conceito geralmente associado à editoração e ao jornalismo. O dicionário Michaelis traz a seguinte definição:

Uso de gráficos e ilustrações para apresentação de informações de forma resumida e atrativa ao leitor. (DICIONÁRIO MICHAELIS, n.d.)

O conceito pode parecer similar ao de visualização de dados, mas existem duas principais diferenças: a presença de narrativas complexas e o uso de ilustrações não envolvendo necessariamente dados. O dataviz, portanto, é uma ferramenta muito importante para a infografia, e jornais que são capazes de se valer de suas estratégias se destacam. Ainda assim, é perfeitamente possível elaborar infográficos importantes sem o uso de dataviz, como no caso do pôster do South China Morning Post (2021) sobre o impacto da pandemia em cidades (figura 11) e no material interativo do New York Times (2019) sobre o incêndio na catedral de Notre Dame de Paris (figura 12).

Figura 11: Versão impressa do South China Morning Post com infográfico sobre mudanças urbanas pós-pandemia. Fonte: South China Morning Post Graphics. Disponível em <<https://twitter.com/SCMPgraphics/status/1380740573196398592>>. Acesso em 29 de novembro de 2021.



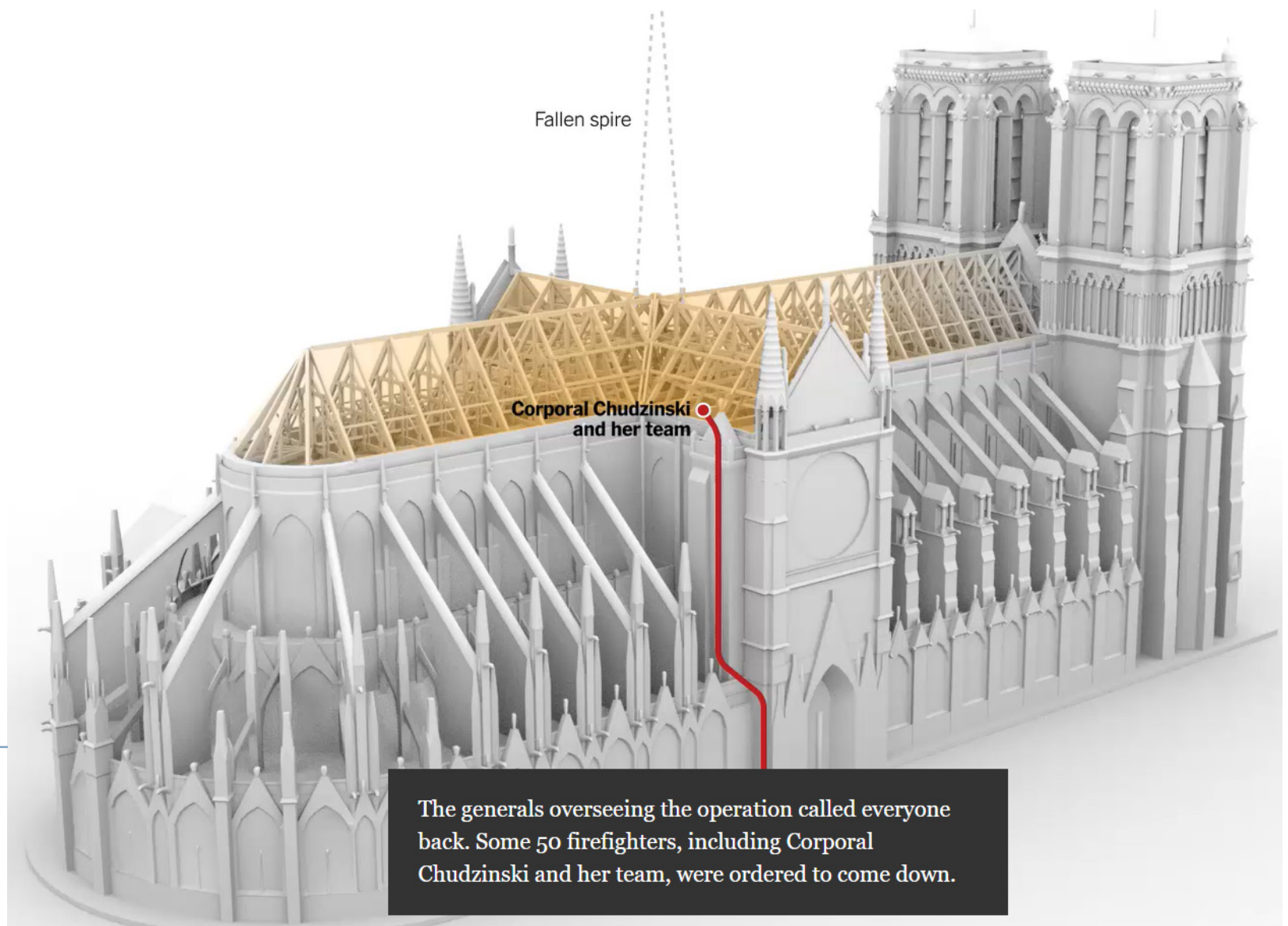


Figura 12: Parte do material interativo sobre o incêndio na Catedral de Notre Dame de Paris em 2019. Fonte: The New York Times. Disponível em <<https://www.nytimes.com/interactive/2019/07/16/world/europe/notre-dame.html>>. Acesso em 29 de novembro de 2021.

Apesar da relação quase inevitável que existe atualmente entre infografia e jornalismo, Sojo (2000) comenta que o termo “infografia” só passou a ser utilizado no jornalismo na década de 1980. Sua primeira utilização se deu em 1963, quando Ivan Sutherland, doutorando do MIT, utilizou um teclado e um lápis óptico para desenhar uma imagem na tela. O prefixo -info estaria, na sua origem, vinculado a informática, não a informação.

O autor ainda cita algumas das primeiras definições do conceito. Em 1985, John Lewell escreveu: “a capacidade de gerar imagens realistas, sombreadas, com cores e completas, com as fontes de luz automaticamente calculadas, o que fará da representação informática, a infografia, uma poderosa rival da fotografia”. Em 1986, Françoise Holtz-Bonneau crava uma definição pragmática para a infografia: “aplicação da informática à representação gráfica do tratamento da imagem” (SOJO, 2000, p. 43).

Esses marcadores temporais são essenciais para alguns autores, e passam a fazer parte da própria definição de infografia. Mesmo referenciando, no mesmo livro, gráficos do século XIX, Moraes (2013) adota uma abordagem sobre infografia baseada na temporalidade. O professor chama a produção anterior à década de 1980 de antecedentes históricos da infografia.

[O reconhecimento da Infografia como sistema], ou seja, a presença do conceito de infográfico assumido na produção jornalística: as peças são feitas segundo o propósito de se esclarecer uma informação para o público por meio da combinação de linguagem verbal e não verbal num único sistema chamado e reconhecido como Infografia. Diferentemente, portanto, do trabalho intuitivo de reunir imagens e palavras numa peça gráfica, a infografia existe a partir do momento em que foi designada para cumprir a referida finalidade de esclarecer e nomeada para tal, o que se deu nos anos 1980. O que foi produzido antes disso, mesmo que para cumprir o papel de explicar, recebeu denominações que variavam de acordo com sua forma: mapa, gráfico de curva, gráfico de barra, perspectiva explodida etc. A partir dos anos 1980, essas categorias foram reunidas (...) sob a denominação Infografia (...). (MORAES, 2013, p.23)

Teixeira (2007) se apega menos ao uso da palavra em si, citando o primeiro infográfico da história da imprensa como tendo sido publicado em Londres, em 1806, sobre o assassinato de um homem. A autora, referenciando De Pablos em artigo para a Revista Fronteiras, cita quatro elementos que são essenciais ao infográfico: um título, um texto de entrada (ou lead, ou linha fina), a indicação das fontes e a assinatura (TEIXEIRA, 2007). A professora propõe uma definição de infográfico a partir de uma indissociável relação entre texto e imagem.

(...) Um infográfico pressupõe a inter-relação indissolúvel entre texto (que vai além de uma simples legenda ou título) e imagem que deve ser mais que uma ilustração de valor exclusivamente estético. Podemos dizer, portanto, que este binômio imagem e texto, na infografia, exerce, por princípio, uma função explicativa e não apenas expositiva. O infográfico, enquanto discurso, deve ser capaz de passar uma informação de sentido completo, favorecendo a compreensão de algo e, neste sentido, nem imagem, nem texto deve se sobressair a ponto de tornar um ou outro indispensável. (TEIXEIRA, 2007, p.113)

Um ponto importante da argumentação da autora é que mapas, tabelas e visualização de dados não são infografia em si. Eles podem ser utilizados como parte de um infográfico, mas a utilização de um gráfico ou uma tabela em uma reportagem jornalística não configura um infográfico (ver figura 13).

A professora usa como exemplo tabelas em certas reportagens: elas servem como ferramentas auxiliares para a compreensão da reportagem, não são elementos interligados.

Teixeira (2007) ainda sugere uma classificação metodológica para tipos de infográficos, dividindo-os em enciclopédicos (explicam temas gerais, como “o que são bactérias”) e específicos (abordam temas específicos e noticiosos, como a explicação de um acidente ou de um resultado eleitoral). Os dois grupos são divididos em infográficos independentes e complementares, em relação à preponderância da infografia no conjunto da matéria. As reportagens infográficas seriam uma subcategoria dos infográficos específicos independentes (TEIXEIRA, 2007).

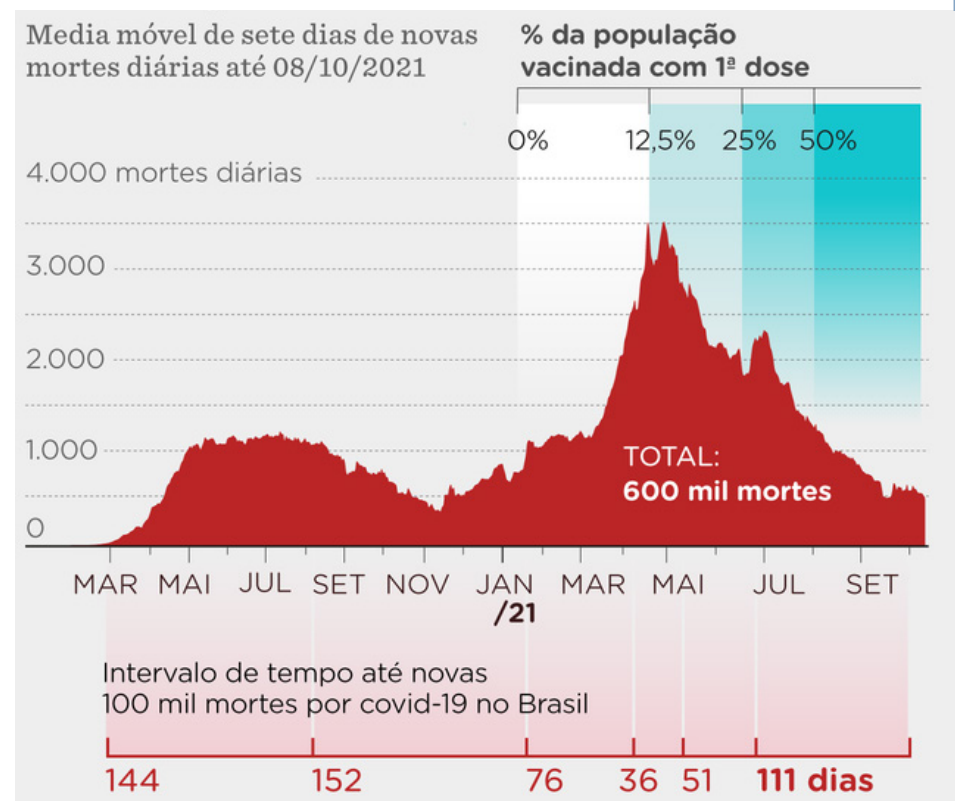
É importante dizer, porém, que a maioria dos autores admite não haver consenso sobre uma definição única de infografia. Em artigo para a Revista Ícone, Teixeira (2010) constata que mesmo em redações não há consenso sobre o que seja infografia. Então, não é de se estranhar que o conceito seja mais ou menos abrangente depen-

do do autor. Mais afeito à visualização de dados, Cairo (2012) argumenta que infografia e visualização existem num contínuo:

Imagine duas linhas retas. Na linha superior, coloco a palavra “Infografia” à esquerda e, na linha inferior, “Visualização” à direita. Todos os gráficos apresentam dados e permitem algum grau de exploração desses dados. (...) Alguns gráficos são quase inteiramente apresentação, permitindo uma exploração limitada; por isso, podemos dizer que são mais infográficos do que visualização, enquanto outros são mais sobre deixar o leitor brincar com o que está sendo mostrado, pendendo para o lado “Visualização” da escala. Mas todo infográfico e toda visualização tem um componente de apresentação e um componente de exploração (...). (CAIRO, 2012)

O que parece caracterizar a infografia, em suma, é a utilização de imagens para estabelecer uma narrativa. Ela se diferenciaria de um livro ilustrado, por exemplo, por sua abordagem jornalística. A visualização de dados pode ser parte do infográfico (ou ele por inteiro), desde que esteja dentro de um contexto narrativo do qual seja indissociável.

Figura 13: Gráfico do jornal Nexo que não seria chamado “infográfico” pela definição de Teixeira (2007), já que conta, sozinho, uma história. Fonte: Nexo Jornal. Disponível em: <https://www.nexojornal.com.br/grafico/2021/10/08/Ap%C3%B3s-20-meses-de-pandemia-Brasil-atinge-a-marca-de-600-mil-mortes>. Acesso em 29 de novembro de 2021.



2.3. Jornalismo de dados

Com a maior disponibilidade de dados, o uso de dataviz em infografia se torna cada vez mais prevalente, mas não apenas nela: os dados passam a fazer parte do cotidiano das redações de jornal, com a exigência de profissionais que sejam capazes não só de representá-los visualmente, mas também de redigir matérias complexas a respeito. Ao uso de análises exploratórias fundamentadas em dados como base para matérias jornalísticas, se dá o nome de “jornalismo de dados” ou “jornalismo guiado por dados”.

Os termos “jornalismo de dados” e “jornalismo guiado por dados” são utilizados como sinônimos, assim como “data journalism” e “data driven journalism”. No Brasil, o termo “jornalismo de dados” é mais popular. Exemplos disso são a Conferência Brasileira de Jornalismo de Dados e Métodos Digitais (Coda) e o Manual de Jornalismo de Dados da Abraji (Associação Brasileira de Jornalismo Investigativo).

Segundo Bounergu (2012), A ideia de “data journalism” tem sua primeira aparição num ensaio de setembro de 2006, A fundamental way newspaper sites need to change. Nele, o programador Adrian Holovaty defende que jornalistas deveriam publicar, além do texto, as fontes dos dados de suas matérias publicamente (BOUNERGU, 2012). É importante salientar, porém, que o termo que Holovaty usa em seu artigo é, na verdade, “journalism via computer programming” (jornalismo

guiado por programação computacional), não “jornalismo de dados”. Após citar uma série de casos “não óbvios” de acontecimentos que poderiam ser analisados pelo jornalismo a partir de sua transformação em bases de dados, o autor critica a resistência das redações em alterar seu modo de pensar e atuar (em tradução própria):

(...) Muitas informações que os jornais coletam são implacavelmente estruturadas. Basta alguém perceber a estrutura (parte fácil) e alguém começar a armazená-la em um formato estruturado (parte difícil). Mas eu consigo entender os jornais serem tão lentos em aceitar esse jeito de pensar. Jornalistas não são o grupo que mais entende de tecnologia, não são a classe mais inovadora e são (só um pouco) resistentes a mudanças. Uma barreira para esse pensamento é um tipo de arrogância jornalística: “Como isso é jornalismo? Nós somos jornalistas, fomos treinados a explicar informações complexas às pessoas de maneira que elas possam entender. Mostrar dados brutos não ajudam pessoas; escrever uma reportagem ajuda as pessoas, porque é inglês simples.” Eu apresentei esses conceitos (“jornalismo guiado por programação computacional”, a importância de dados legíveis por máquinas etc.) em um

número de eventos da indústria jornalística, e inevitavelmente alguém pergunta essas questões. (HOLOVATY, 2006)

A proposta de utilizar bases de dados para matérias jornalísticas, porém, é anterior ao século XXI. Träsel (2014) inicia sua história do jornalismo de dados falando sobre jornalismo de precisão. Em 1967, o repórter Philip Meyer decidiu cobrir agitações sociais que aconteciam em Detroit de uma maneira pouco habitual: com o apoio do psicólogo Nathan Caplan e de uma equipe de 30 entrevistadores, obteve uma amostra representativa dos residentes dos bairros afetados. Com esses resultados, descobriu-se que:

(...) a) não havia correlação entre condição econômica e participação nos distúrbios: b) nativos da cidade eram três vezes mais propensos a ter participado de saques, agressões, homicídios e incêndios do que os imigrantes do Sul. Além disso, descobriu-se que as razões principais para o motim foram a brutalidade policial, as más condições de habitação e a escassez de empregos. Por outro lado, o relatório também mostrou que os participantes dos distúrbios eram um subgrupo específico que não representava a atitude geral da comunidade negra. (TRÄSEL, 2014, p.97)

A partir dessa experiência, Meyer publicou um livro chamado *Precision Journalism*. Nele, falava de metodologias que havia aprendido na faculdade de ciências sociais envolvendo computação, que poderiam ser utilizadas por jornalistas. Apesar da importância do computador, o grande diferencial está nos métodos de pesquisa aplicados: um repórter poderia perfeitamente fazer jornalismo de precisão, portanto, sem o auxílio de computadores.

Träsel argumenta que a principal divergência teórica do jornalismo de precisão era com o Novo Jornalismo, prática difundida nos EUA nos anos 1960 que buscava trazer elementos literários à prática jornalística; o interesse, portanto, seria de trazer a objetividade de volta às redações. (TRÄSEL, 2014)

O conceito de jornalismo de precisão não perdurou, tendo se popularizado a expressão “reportagem assistida por computador” para definir prática semelhante, mas com maior ênfase na utilização da informática do que na metodologia científica envolvida. Apesar disso, a ideia remete à prática contemporânea de jornalismo de dados.

Se o jornalismo de precisão podia ser dissociado da informática, numa realidade em que grandes bases de dados precisam ser analisadas, o uso da computação se torna imprescindível. A grande mudança paradigmática do jornalismo de dados atual, portanto, está justamente na necessidade de compreender e divulgar a quantidade de informação disponível hoje.

No *Data Journalism Handbook*, Bounergu (2012), logo após abordar possíveis diferenças metodológicas entre “reportagem assistida por computador” e “jornalismo de dados”, adota uma postura que reforça a importância do contexto histórico para a prática, sem menosprezar os desenvolvimentos anteriores (em tradução própria):

(...) Alguns argumentam não haver diferença significativa entre jornalismo de dados e reportagem assistida por computador. É senso comum que mesmo as práticas de mídia mais recentes apresentam narrativa, assim como algo novo nelas. Ao invés de debater se jornalismo de dados é ou não completa novidade, uma posição mais frutífera seria considerá-lo parte de uma tradição mais longa, mas respondendo a novas circunstâncias e condições. Mesmo que possivelmente não existam diferenças de objetivos e métodos, a emergência do rótulo “jornalismo de dados” no começo do século indica uma nova fase em que o imenso

volume de dados que estão disponíveis online — junto com ferramentas focadas no usuário mais sofisticadas, autopublicação e ferramentas de crowdsourcing — permitem que mais pessoas trabalhem com mais dados mais facilmente do que nunca. (BOUNERGU, 2012, p.21)

No mesmo livro, Bradshaw (2012) reforça a importância do contexto histórico para o desenvolvimento da área (em tradução própria):

O que é jornalismo de dados? (...) Tanto “dados” quanto “jornalismo” são termos problemáticos. Algumas pessoas pensam em “dados” enquanto uma coleção de números, provavelmente agrupados numa planilha. 20 anos atrás, isso era basicamente a única fonte de dados com a qual jornalistas tinham que lidar. Mas nós vivemos num mundo digital, um mundo em que quase tudo pode (e quase tudo é) descrito com números. (...) O que torna o jornalismo de dados diferente do resto do jornalismo? Talvez sejam as novas possibilidades que se abrem quando você combina o tradicional “faro por notícias” e a habilidade de contar uma história envolvente com a escala e variedade da informação digital hoje disponível. (BRADSHAW, 2012, p.2)

A dimensão histórica também é importante para Coddington (2014), mas o autor dá ênfase ao rompimento com o jornalismo investigativo e no uso da visualização de dados:

Embora jornalismo de dados seja um termo geralmente no contexto de projetos investigativos (...), é menos atrelado ao jornalismo investigativo do que a RAC (Reportagem Assistida por Computador). (...) Outras distinções incluem a ênfase do jornalismo de dados na visualização como prática fundamental, havendo uma profunda relação entre o design da visualização e os valores jornalísticos, e um rompimento epistemológico em que o jornalismo de dados vê os leitores como co-construtores de verdades e valores morais. (CODDINGTON, 2014, p.335)

É interessante acrescentar que o autor ainda referencia uma terceira terminologia, menos popular entre as redações: “jornalismo computacional”. Coddington cita definições acadêmicas que buscam abarcar tanto a RAC quanto o jornalismo de dados na mesma definição, mas entende que esse tipo de caracterização dificulta a diferenciação entre as três práticas. Assim, cria sua própria definição de jornalismo computacional:

Uma vertente do jornalismo orientado por tecnologia centrado na aplicação de computação e pensamento computacional

para as práticas de coleta de informações, “sense-making” [processo de dar sentido à informação] e apresentação da informação, ao invés do uso jornalístico de dados ou de métodos de ciências sociais”. (CODDINGTON, 2014, p.335)

Apesar da ênfase das definições nos datasets que podem ser utilizados, é importante diferenciar o trabalho de um jornalista de dados para o de um cientista de dados tradicional. Além de o objetivo e o consumidor final das análises da imprensa serem distintos em relação aos de uma empresa ou instituição, o resultado do trabalho é fundamentalmente diferente. O’Murchu (2012) explicita isso, defendendo as práticas do jornalismo de dados como essenciais à profissão (em tradução própria):

Acho que é importante destacar o “jornalismo” em “jornalismo de dados”. O ofício deve ser não apenas sobre analisar ou visualizar dados “porque sim”, mas para usar como uma ferramenta para chegar mais perto da verdade do que está acontecendo no mundo. Vejo a habilidade de analisar e interpretar dados como uma parte essencial do jornalismo contemporâneo, ao invés de uma disciplina separada. No fim, tudo é sobre boas reportagens e sobre contar histórias da maneira mais apropriada. (O’MURCHU, 2012, p. 9)

Vasconcellos, Mancini e Bittencourt (2015) buscam criar uma categorização para diferenciar “jornalismo com dados” e “jornalismo de dados” no Brasil. Os autores buscam trabalhar com três dimensões que, quanto mais presentes, mais permitem chamar uma matéria de “jornalismo de dados” (ver figura 14).

Eles englobam a dimensão investigativa, ao identificar ou não se a reportagem revisita o processo de extrair e organizar o material bruto; a dimensão interpretativa, verificando se existe um texto jornalístico analítico; e uma dimensão comunicativa, relacionada à presença de gráficos e infográficos (VASCONCELLOS et al, 2015). Os autores acabam usando “visualização de dados” e “infografia” como sinônimos, premissa contrária à definição de infografia por Teixeira (2007), apresentada no capítulo anterior.

A diferenciação proposta para discernir jornalismo com dados de jornalismo de dados lembra a definição de infografia dada por Teixeira (2007), que diferencia “gráfico” e “infográfico” justamente pelo grau de interligação entre a ilustração e a reportagem. Nota-se, dessa forma, a busca por uma ideia de infográfico/jornalismo de dados “ideal” como sendo pautado pela investigação proposta pela prática jornalística.

Tanto teóricos da infografia quanto do jornalismo de dados parecem compreender a outra área como uma parcela do seu campo de atuação. Para Teixeira (2007), as metodologias de jornalismo de dados podem contribuir para a realização de um infográfico enciclopédico; para Vasconcellos et al (2015), a infografia (usada como sinônimo de visualização de dados) é um dos componentes que caracterizariam uma reportagem de jornalismo de dados

“nível 1”. Isso não significa que os autores acreditem que a outra área seja uma sub-área do seu campo de estudo (em nenhum momento há algum juízo de valor), mas mostra como essas áreas acabam se cruzando na teoria e na prática. A visualização de dados, em contraste, é entendida como uma ferramenta para os dois campos.

Mesmo com tantas definições, é importante salientar que, assim como a visualização de dados e a infografia, o jornalismo de dados está em constante transformação. Um fator importantíssimo é a ênfase crescente no jornal online: as estratégias de visualização e de disseminação da informação são muito diferentes hoje do que eram há cinco anos. Se alguém de fora do jornalismo produz um texto investigativo sobre sua área de atuação, com abordagem baseada em dados, pode dizer que está fazendo jornalismo de dados? Há perguntas que só observando a prática jornalística nos próximos anos conseguiremos responder.

Categorias X Níveis de JD		Busca e/ou Elaboração própria dos dados (criação da base)	Estrutura a base (séries temporais, categorias rankings, tabelas)	Visualização dos dados (infografia)	Interpretação dos dados (texto)
Jornalismo → Com Dados	Nível 1				
	Nível 2				
	Nível 3				
	Nível 4				
	Nível 5				

Figura 14: Tabela proposta para mensurar reportagens numa escala de “totalmente jornalismo com dados” até “totalmente jornalismo de dados”. Fonte: Vasconcellos et al, 2015, p.17.

3. Métodos

3.1. TFG 1 (agosto a dezembro de 2021)

Desde o início da disciplina de TFG 1, em agosto de 2021, já tinha clareza de que queria fazer um trabalho envolvendo visualização de dados.

Dois trabalhos que me fizeram começar a pensar em possibilidades dentro de um trabalho de conclusão de curso em Arquitetura e Urbanismo foram os de Eric Palmeira (2020), que fez uma plataforma para visualizar a idade dos edifícios do município de São Paulo, e de Gabriela Momberg Araújo (2021), que propôs um sistema, em estampa, de visualizações de dados sobre a qualidade da alimentação no município.

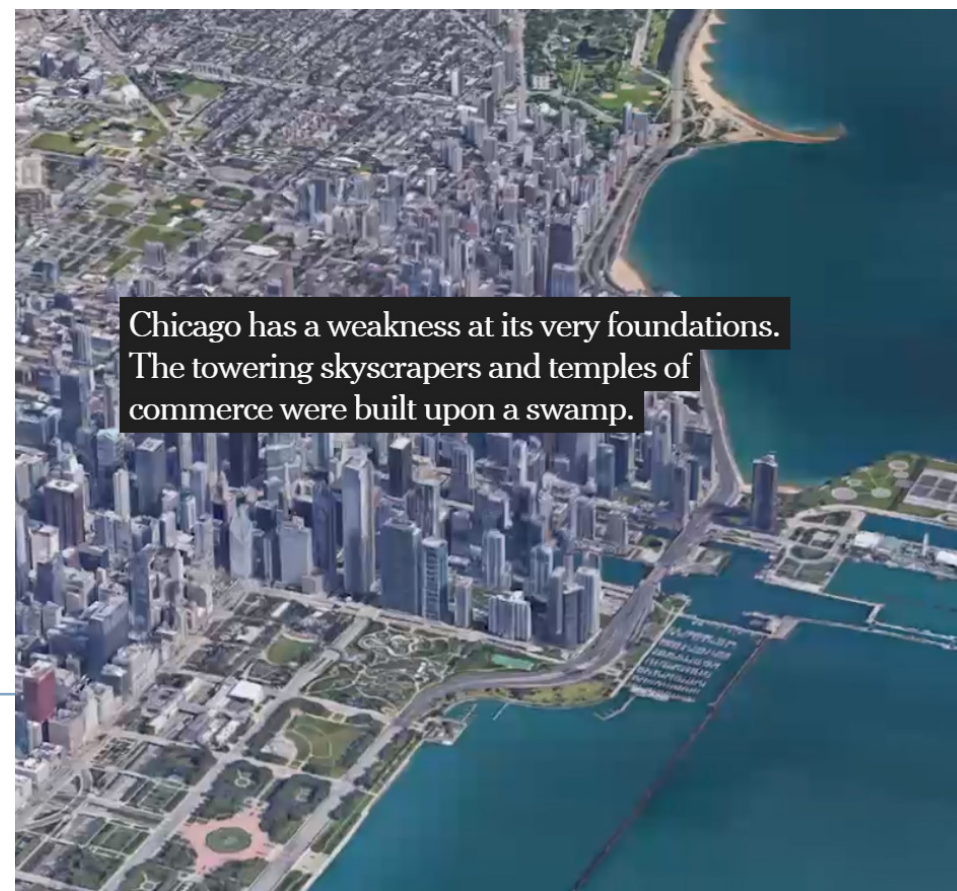
Decidi explorar a visualização de dados a partir de uma narrativa, proposta evidentemente influenciada por estar trabalhando num jornal, à época.

O Prof. Leandro Velloso, orientador deste trabalho, comentou que eu deveria buscar algum tema estimulante, que tivesse alguma relevância jornalística (esse era meu arcabouço) e fosse exequível: não é possível fazer um trabalho em dados se o principal (os dados) não existem, são incompletos ou de muito difícil acesso.

Com essa tarefa em mente, me deparei com a matéria “A Battle Between a Great City and a Great Lake”, do New York Times. Nela, Dan Egan discorre sobre a relação histórica que Chicago tem com o Lago Michigan e com a bacia do rio Mississipi, além de como as mudanças climáticas podem afetar o próprio futuro da cidade (EGAN, 2021).

A reportagem não é exemplo de jornalismo de dados, mas tem uma narrativa envolvente e apresenta qualidades infográficas ao utilizar o modelo 3D de Chicago para localizar alguns locais fundamentais à história (figura 15).

Figura 15: Trecho da reportagem sobre Chicago e a emergência climática. Fonte: The New York Times. Disponível em: <https://www.nytimes.com/interactive/2021/07/07/climate/chicago-river-lake-michigan.html>. Acesso em 19 de novembro de 2021.



Com a reportagem sobre Chicago, me dei conta de que já havia tido muito contato com matérias sobre a emergência climática, mas que elas costumavam ser trabalhadas numa escala mais ampla do que a municipal.

Decidi estudar os impactos das mudanças climáticas em áreas urbanas. Conversei com a Prof.^a Denise Duarte sobre isso e ela aceitou ser co-orientadora do trabalho, ajudando a construir uma base mais sólida sobre o tema. Foi a Prof.^a, também, que falou sobre a existência de Planos de Ação Climática em alguns lugares, documentos que contextualizam a crise no âmbito municipal e propõem medidas de mitigação e adaptação locais.

O principal material necessário para se fazer visualização de dados são, justamente, os dados. Por isso, ao pensar em uma análise a nível municipal, uma barreira importante é a disponibilidade de bases de dados públicas. A abordagem de grandes metrópoles não seria somente por conta de sua relevância, mas também por serem locais em que o nível de debate público estimulou que muitas informações fossem disponibilizadas online gratuitamente.

Inicialmente, este trabalho seria focado em várias cidades da América Latina. A escolha da região de estudo traria uma perspectiva do Sul global ao tema, mas também limitaria os idiomas das bases de dados a Português e Espanhol, um aspecto prático muito relevante.

O primeiro passo do trabalho, portanto, foi averiguar quais metrópoles latino-americanas possuíam algum site com bases de dados públicas. Ao pegar a lista de dez maiores cidades das Américas, Nova York e Los An-

geles foram descartadas por razões geográficas. Lima, Santiago e Caracas não tinham portais de dados públicos de fácil acesso. Das restantes, Buenos Aires, Cidade do México e São Paulo pareciam ter muita informação disponível, além de Planos de Ação Climática.

O trabalho inicialmente trataria da emergência climática nas três cidades. Todas as bases de dados encontradas foram tabuladas e ranqueadas conforme a pertinência ao tema, definida por checagem visual (ver figura 16).

Bases de dados TFG

Arquivo Editar Ver Inserir Formatar Dados Ferramentas Extensões Ajuda A última edição foi feita em 19 de abril

100% RS % .00 123 Calibri 11 B I A

Interesse					
A	B	C	D	E	F
Interesse	Nome	Descrição	Links		
1	Calidad de Aire	Información diaria de los niveles medios de monóxido de carbono, dióxido de nitrógeno y material particulado respirable menor a 10 micrones medidos en las estaciones de La Boca, Córdoba, Palermo...	osaires.gob.ar/dataset/calidad-aire		
1	Bicicletas Públicas	Origen, destino, horario, género y edad de los viajes realizados en el sistema de bicicletas públicas de la Ciudad	ires.gob.ar/dataset/bicicletas-publicas		
6	Premetro	Ubicación geográfica de las líneas y estaciones del premetro.	nosaires.gob.ar/dataset/premetro		
7	Bicicleteros en Vía Pública	Listado con ubicación geográfica de los bicicleteros instalados en la Ciudad.	s.gob.ar/dataset/bicicleteros-via-publica		
8	Cicloviás	Referencia geográfica de las cicloviás de la Ciudad	nosaires.gob.ar/dataset/ciclovias		
10	Metrobus	Estaciones y recorrido con ubicación geográfica del Metrobus.	nosaires.gob.ar/dataset/metrobus		
11	Bicicleterías	Listado con ubicación geográfica de bicicleterías de la Ciudad.	osaires.gob.ar/dataset/bicicleterias		
12	Espacios Verdes	Límites y ubicación geográfica de los espacios verdes de la Ciudad (jardín, parque, patio recreativo, plaza, plazoleta, cantero y polideportivo).	aires.gob.ar/dataset/espacios-verdes		
15	Unidades Territoriales de Inclusión Urbana	Información sobre las Unidades Territoriales de Inclusión Urbana (UTIU) cuya delimitación geográfica se hizo considerando los aspectos físicos y sociales, para proponer áreas de oportunidad a...	dataset/unidades-territoriales-inclusion-urbana		
18	Viajes en Transporte Público	Cantidad de viajes diarios realizados en transporte público de la ciudad según modo de transporte. Información disponible en la API de Transporte de la Ciudad.	enosaires.gob.ar/dataset/sube		
19	Flujo vehicular	Cantidad de vehículos detectados en los sensores de tránsito de la Ciudad. Información disponible en la API de Transporte de la Ciudad.	zob.ar/dataset/flujo-vehicular-anillo-digital		
14	Grandes Contribuyentes	Información de los grandes contribuyentes definidos por la Administración Gubernamental de Ingresos Públicos (AGIP).	s.gob.ar/dataset/grandes-contribuyentes		
15	Alto Riesgo Fiscal	Padrón de alto riesgo fiscal para los contribuyentes responsables del Impuesto sobre los Ingresos Brutos.	aires.gob.ar/dataset/alto-riesgo-fiscal		
16	Subte: Viajes Molinetes	Cantidad de pasajeros por molinete y por estación de todas las estaciones de la red de subte.	s.gob.ar/dataset/subte-viajes-molinetes		

Figura 16: Parte de planilha com os datasets encontrados para Buenos Aires, obtidos por raspagem do site Buenos Aires Data. Fonte: o autor.

3.2. TFG 2 (março a julho de 2022)

O planejamento previsto para o segundo semestre consistia no fluxo de trabalho com o qual já estava familiarizado pelo estágio no jornal. Uma leitura inicial de bibliografia para dar suporte ao assunto; a exploração das bases de dados disponíveis; a produção de gráficos sem tratamento final para conseguir colocá-los em uma narrativa, concomitante à escrita do roteiro; e, finalmente, a elaboração de infográficos.

A realidade se mostrou mais complexa do que isso. Apesar de ter reservado três semanas do semestre para a leitura inicial, “emergência climática em espaços urbanos” se mostrou tema amplo demais para o tempo disponível.

Em abril de 2022, tive uma conversa informal com Rodolfo Almeida, infografista e pesquisador da UFRJ, sobre dataviz e mudanças climáticas. Rodolfo me alertou que o tema realmente pode ser paralisante e complexo demais, e me sugeriu explorar as bases de dados que havia coletado o quanto antes.

Buenos Aires foi a primeira cidade que decidi explorar, por ordem alfabética. A leitura do Plano de Ação Climática da cidade foi feita simultaneamente à exploração das bases de dados catalogadas como de “interesse 1” no semestre anterior.

Os datasets se mostraram mais difíceis de trabalhar do que o esperado, com informações faltantes ou de difícil compreensão para quem não está familiarizado com a dinâmica da capital argentina. Ter o Plano de Ação Climática como principal referência também me fez temer adotar uma narrativa acrítica dos fatos, já que uma prefeitura vai contar sua versão e destacar os feitos daquele governo.

Conversando com a Prof.^a Denise Duarte, decidi limitar o trabalho a São Paulo. A escolha se deu, principalmente, por dois fatores: a facilidade de obter informações além do geportal oficial da prefeitura (Geosampa), já que seria possível consultar a própria professora sobre bases de dados disponíveis; e a possibilidade de encontrar bibliografias críticas à versão oficial dos fatos com mais facilidade.

A água é um elemento presente em várias como fator sensível às alterações climáticas, especialmente em São Paulo, então foi uma escolha natural para objeto de foco. Ela também é presente na principal referência de narrativa (a reportagem sobre Chicago).

A elaboração do caderno que acompanha este relatório foi um processo concomitante de escrita de roteiro e de exploração de bases de dados, além da própria confecção dos gráficos. Para conseguir pensar numa narrativa coerente, ela foi dividida em

quatro partes. O caderno não apresenta divisões, mas essa estrutura facilitou pensar em quais recursos precisariam ser investigados para dar conta de cada tópico.

Os quatro tópicos da narrativa:

1. O histórico água-cidade;
2. Os processos de desequilíbrio;
3. A emergência climática e seus impactos;
4. As alternativas e o que pode ser feito.

A metodologia de cada infográfico será apresentada no capítulo 7 deste relatório. Ele também inclui discussões sobre os diversos formatos utilizados.

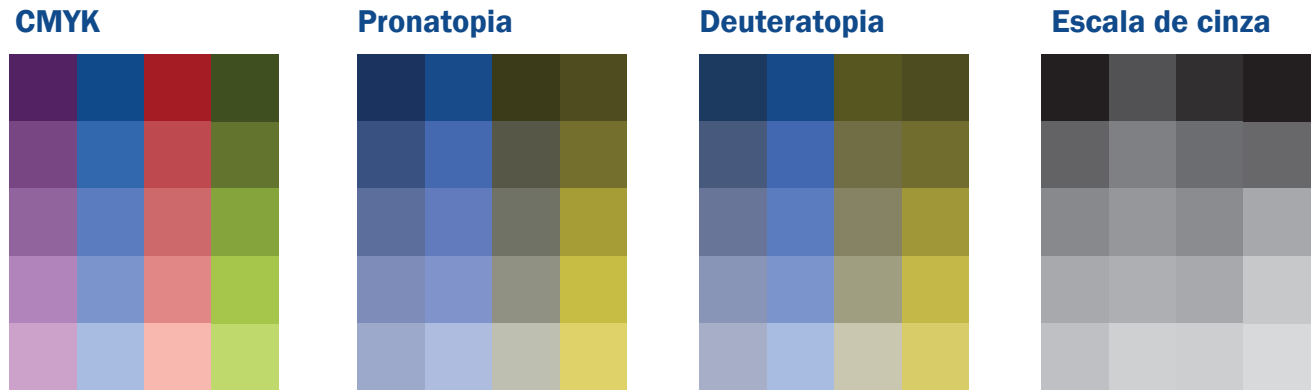
3.3. Guia de Estilo

A aparência é aspecto fundamental do trabalho com visualização de dados. Não basta um gráfico ser informativo: ele deve prender a atenção do leitor. No caso de um caderno que reúne vários infográficos, deve também apresentar coerência visual.

Em relação à paleta de cores, a inspiração veio daquele que é possivelmente o mais famoso gráfico recente sobre mudanças climáticas. As listras climáticas (climate stripes, também chamadas warming stripes), que serão discutidas no capítulo 7, utilizam um degradê do azul escuro ao vermelho escuro para representar a mudança na temperatura média.

A paleta deste trabalho é inicialmente composta pelos azuis e vermelhos mais escuros e mais claros do gráfico de listras climáticas, além de três tons intermediários. A mistura de azul e vermelho (roxo) foi escolhida como uma terceira cor básica. O verde foi escolhido como última cor da paleta, pois a cartografia costuma utilizar a cor para vegetação, mas numa família mais amarelada, para evitar problemas de acessibilidade. As cores passaram pelos testes de acessibilidade em pronatopia e deuteratopia (formas de daltonismo) e em escala de cinza. Os infográficos também foram todos analisados utilizando esses filtros do Illustrator.

Cores adotadas no caderno



Mais para
cima: maior
prioridade
de uso

C 100 M 80 Y 18 K 3	C 86 M 70 Y 4 K 0	C 68 M 48 Y 0 K 0	C 36 M 56 Y 0 K 0	C 32 M 20 Y 0 K 0
C 23 M 100 Y 95 K 19	C 19 M 84 Y 66 K 8	C 15 M 50 Y 70 K 4	C 8 M 55 Y 38 K 1	C 0 M 23 Y 33 K 0
C 78 M 100 Y 28 K 20	C 62 M 85 Y 19 K 5	C 48 M 69 Y 9 K 0	C 31 M 52 Y 0 K 0	C 17 M 40 Y 0 K 0
C 67 M 45 Y 100 K 45	C 60 M 35 Y 100 K 23	C 53 M 18 Y 100 K 3	C 40 M 6 Y 91 K 0	C 29 M 0 Y 74 K 0

Como a inspiração inicial para o caderno foi o mundo jornalístico, para a tipografia, analisou-se aquelas utilizadas em alguns jornais. Havia fontes diferentes para os títulos das matérias, os textos e as linhas finas (trecho de texto acompanha o título). A maioria dos veículos de comunicação seguiam regras diferentes para as matérias e para os gráficos (ver figura 17).

Para pensar o caderno, não fazia sentido adotar fontes para “título” e “linha fina”, já que ele é composto por páginas de texto corrido e páginas de infográficos. Em todos os veículos analisados, a fonte de corpo do texto é serifada e, em quase todos, a fonte dos gráficos não é. Decidiu-se, portanto, por uma fonte serifada para o corpo de texto (Merriweather) e uma fonte sem serifa para todos os infográficos (Franklin Gothic).

Pela ordem de prioridade cromática (figura 19), o azul mais escuro da paleta foi a cor escolhida para o texto. Por conta de contraste, alguns textos de legenda de infográficos utilizam branco.

Quanto ao formato, foi utilizada como referência a obra “México, Ciudad Futura”, de Alberto Kalach (2010), cujas páginas têm 18,5x22,5 cm.

Veículo	Título	Linha Fina	Corpo do texto	Gráficos
The New York Times	Cheltenham Std (serifada)	Não identificada (serifada)	Georgia (serifada)	Franklin Gothic (não-serifada)
Folha de São Paulo	Folha II Bold (serifada)	Não identificada (não-serifada)	Folha Texto (não-serifada)	Folha Gráfico (não-serifada)
La Nación	SuecaSlab (serifada)	Arial Regular (não-serifada)	Georgia (serifada)	Não identificada (serifada)
Le Monde	The Antigua B (serifada)	Franklin Gothic (não-serifada)	Georgia (serifada)	Franklin Gothic (não-serifada)
Financial Times	Financier (serifada)	MetricWeb (não-serifada)	Georgia (serifada)	Benton Sans (não-serifada)

Figura 17: Tabela com as tipografias de jornais selecionados. Fonte: o autor.

Tipografia adotada no caderno

Merriweather Regular para texto normal

Merriweather Black para destaques

Franklin Gothic Demi para infográficos

Franklin Gothic Book para legendas

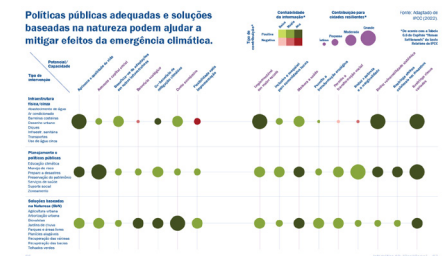
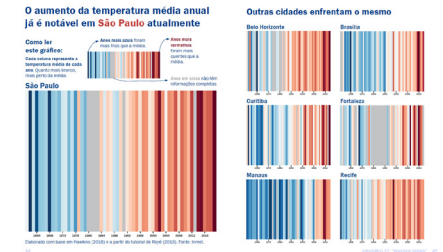
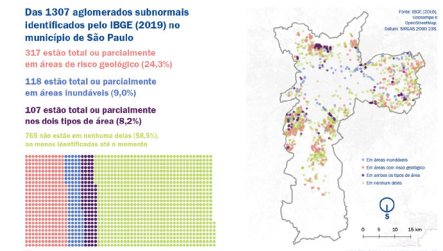
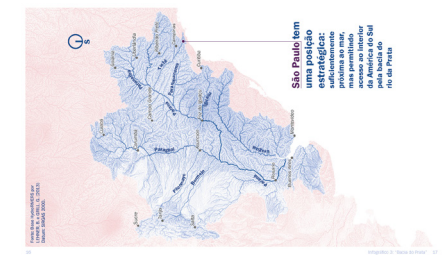
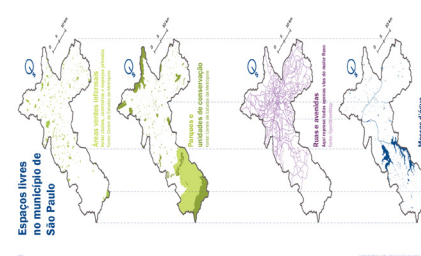
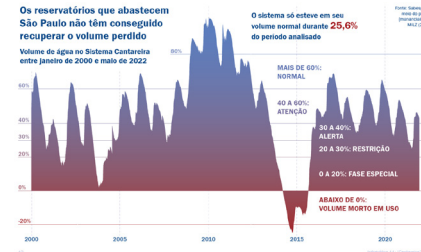
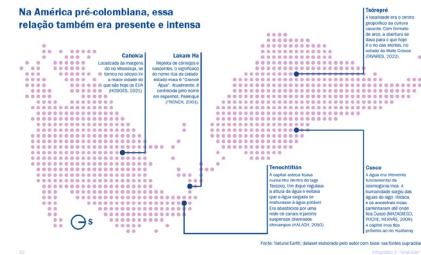
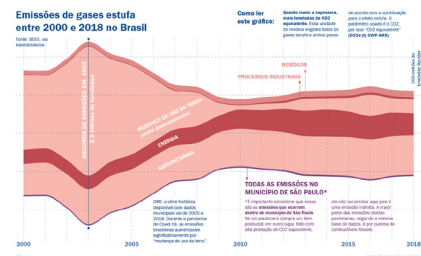
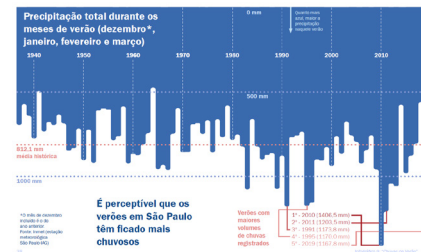
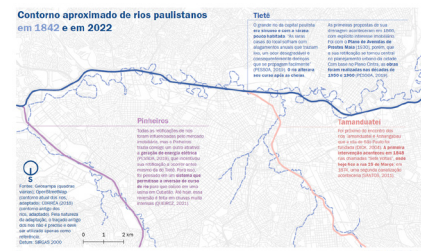
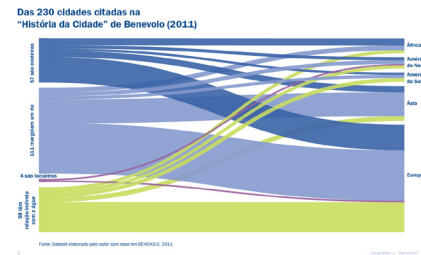
4. Fichas dos infográficos

A elaboração de cada infográfico seguiu uma linha de produção comum:

1. Identificar bases de dados adequadas à narrativa;
2. Se necessário, tratá-las utilizando Excel ou R;
3. Elaborar “gráficos crus”, ou seja, primeiras versões sem acabamento visual elaborado, em QGIS ou R (pacote ggplot2);
4. Tratar visualmente utilizando QGIS (mapas) e/ou Illustrator (mapas e gráficos);
5. Realizar testes de acessibilidade com filtros do Illustrator;
6. Adicionar legendas faltantes no Indesign.

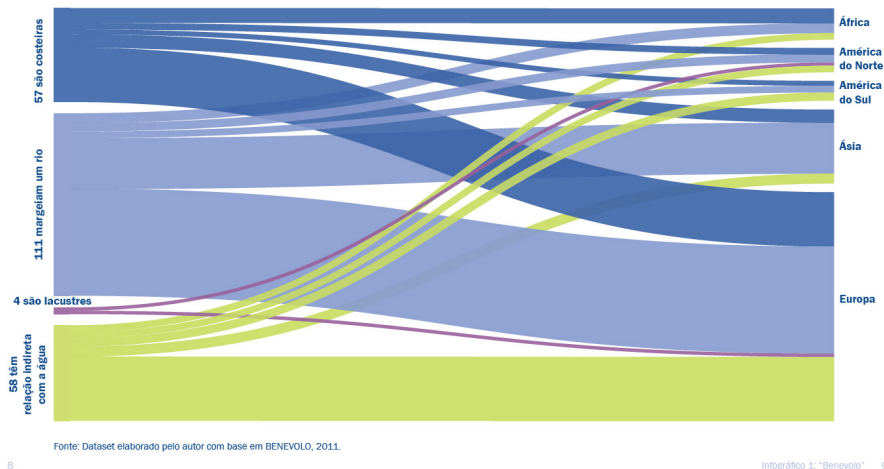
As bases de dados públicas estão disponibilizadas num repositório no Github. O acesso para ele é em <<https://github.com/nicholaspretto/TFG>>.

Apesar dos pontos em comum, cada gráfico teve um processo distinto de elaboração, explicado nos tópicos a seguir. A ênfase da discussão pode se dar no formato do gráfico ou no processo de obtenção e tratamento dos dados, dependendo do que tenha sido identificado como mais marcante ou pertinente.



Infográfico 1: Benevolo

Das 230 cidades citadas na
“História da Cidade” de Benevolo (2011)



Descrição: Diagrama de fluxo do tipo Sankey mostrando o número de cidades descritas no livro “História da Cidade” de Leonardo Benevolo. A relação descrita é o número de cidades de acordo com sua relação com a água, à esquerda, e os continentes em que elas se encontram, à direita.

Objetivo: Iniciar a narrativa mostrando a importância da água na história das cidades de modo amplo. Utilizar um gráfico que, com seu formato, remeta à fluidez da água.

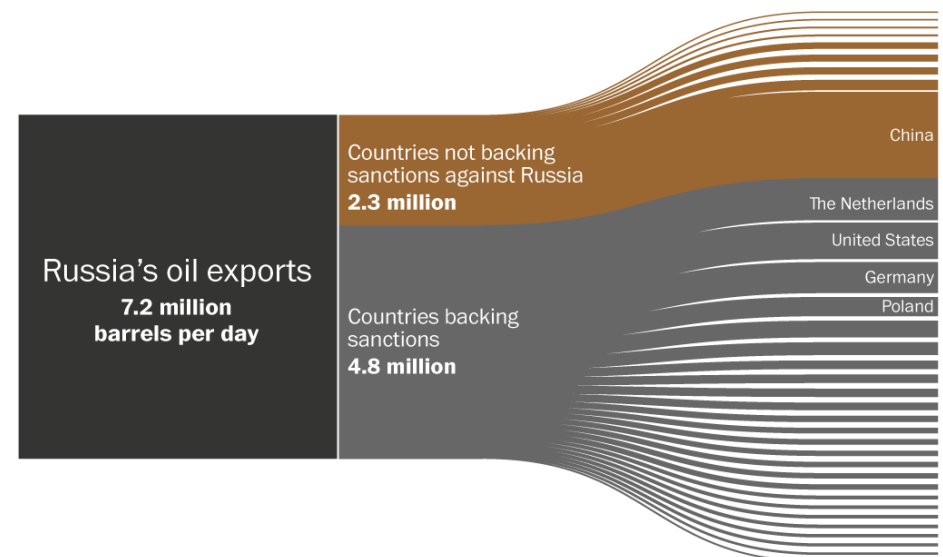
Fonte: Dataset próprio elaborado a partir do livro “História da Cidade” de Leonardo Benevolo (2011).

Metodologia: Foi elaborada uma revisão do livro citado com o auxílio do Excel, incluindo o nome das localidades e sua relação com a água. A determinação do status ocorreu por meio da descrição no próprio livro ou, quando a informação estivesse indisponível, por análise a partir do Google Maps. A partir da categorização, utilizou-se a função “Alluvial Diagram” do site RAWGraphs para o layout básico e o Illustrator para finalização.

Discussão: O gráfico do tipo Sankey visa mostrar uma informação quantitativa, geralmente de alto valor, distribuída em diferentes categorias. Seu uso costuma ser associado a um diagrama de fluxo, já que a espessura das linhas pode variar ao longo de um processo.

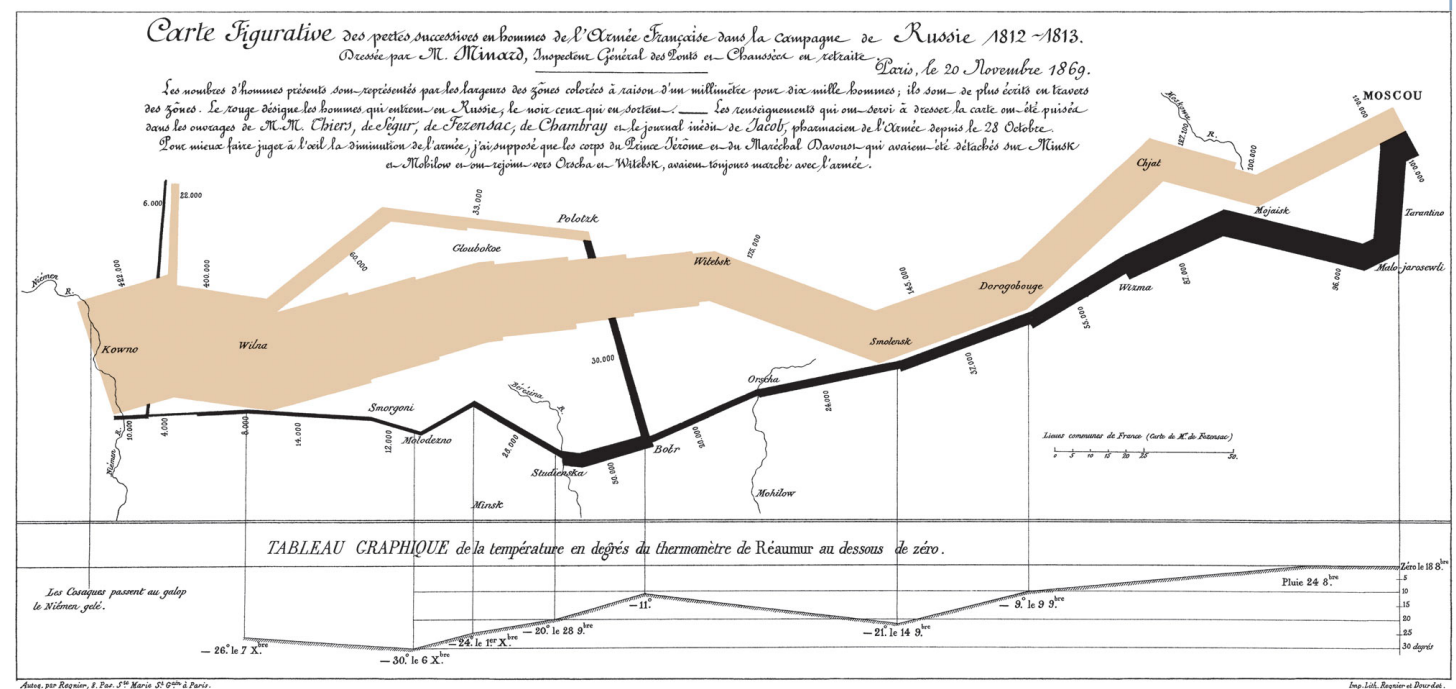
Segundo Phineas (2007), ele tem esse nome por conta de um gráfico de Matthew Sankey, capitão da marinha britânica no século XIX, mostrando a eficiência de um motor a vapor. Nele, a espessura das linhas são relativas à perda de calor em cada etapa ilustrada. Os diagramas de Sankey são muito utilizados na ciência e na engenharia. Variações mais simples dele, sem a ideia de ciclo, são recorrentemente utilizadas no jornalismo (figura 18).

Figura 18: Diagrama tipo Sankey mostrando o destino das exportações de petróleo da Rússia. Fonte: Washington Post. Disponível em <https://www.washingtonpost.com/business/2022/03/08/russia-oil-imports-ban/?utm_source=twitter&utm_campaign=wp_graphics>. Acesso em 30 de maio de 2022.



A fama se deve, principalmente, a Edward Tufte descrever o gráfico como “provavelmente o melhor gráfico estatístico já produzido” (TUFTE, 2007, p.40-41). O diagrama de Minard ainda é muito admirado e referenciado (Figura 20).

Figura 19: Diagrama de fluxo proposto por Charles Minard. Fonte: TUFTE, 2007.



Which E.U. countries import the most Russian gas

Share of country's natural gas imports from Russia, 2020

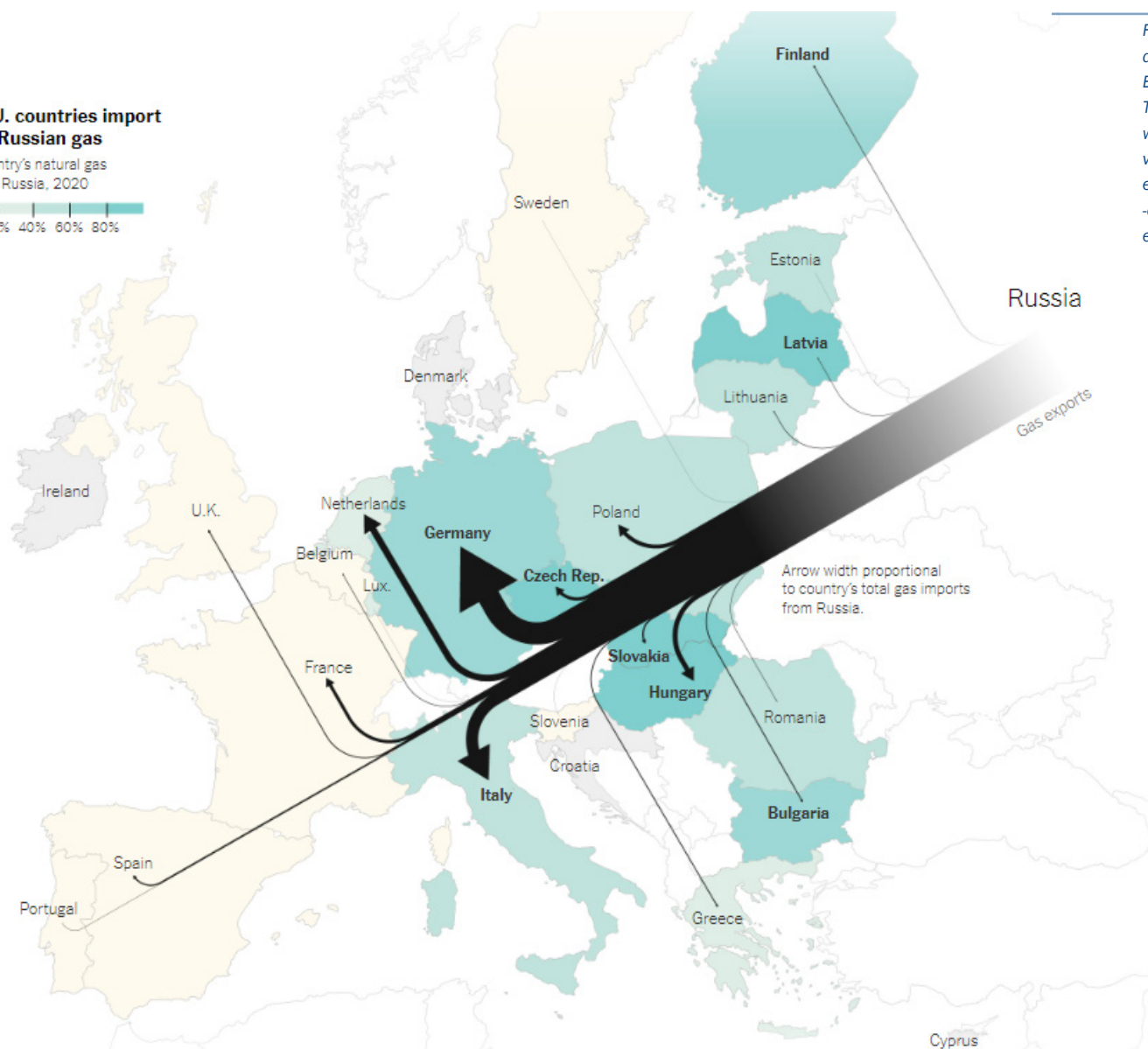
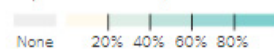
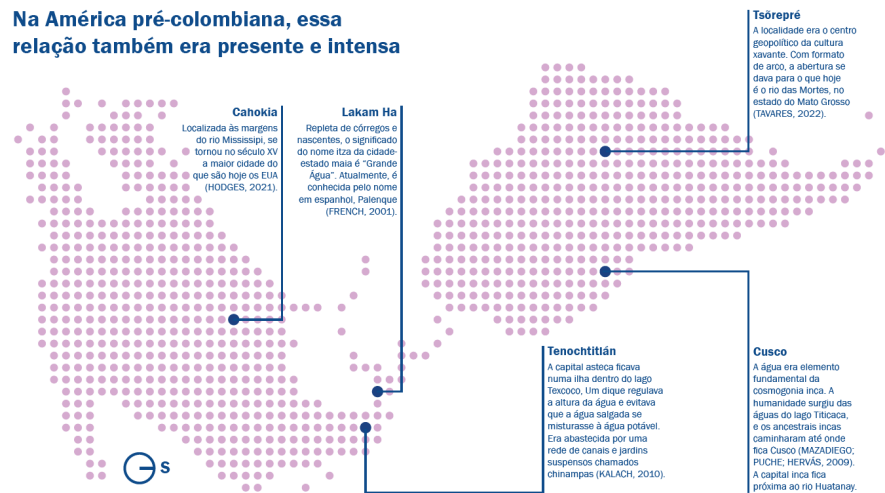


Figura 20: Mapas com os fluxos de gás natural russo para a Europa. Fonte: The New York Times. Disponível em <<https://www.nytimes.com/interactive/2022/02/15/business/energy-environment/russia-gas-europe-ukraine.html>>. Acesso em 30 de maio de 2022.

Infográfico 2: Américas

Na América pré-colombiana, essa relação também era presente e intensa



Fonte: Natural Earth; dataset elaborado pelo autor com base nas fontes supracitadas.

Infográfico 2: "Américas" 11

Descrição: Mapa estilizado das Américas com o norte para a esquerda, com pontos representando o continente. Pontos destacados apresentam exemplos de agrupamentos humanos com desenvolvimento anterior à colonização europeia.

Objetivo: Apresentar a ancestralidade da relação cidade-água sob uma perspectiva não-europeia, destacando os povos originários das Américas. Explorar técnicas de cartografia estilizada.

Fonte: Natural Earth para a base cartográfica. Múltiplas referências para os textos de cada localidade, descritas no caderno.

Metodologia: Com o mapa-múndi do projeto Natural Earth, foi gerada por meio do QGIS uma base de pontos do continente americano. Os pontos destacados foram coloridos no Illustrator.

Discussão: Pensar em figuras a partir de unidades menores é a base de várias estratégias para reprodução de imagens em telas: cada pixel é um pequeno quadrado monocromático que, no conjunto, compõe uma imagem completa. A palavra "raster" é utilizada para descrever essas imagens, em contraste com arquivos vetorizados.

Essa lógica encontra precedente no Pontilhismo, movimento artístico que surgiu na França na segunda metade do século XIX (figura 21). O estilo tem maior ênfase na técnica do que em temas ou em abordagens conceituais específicas, propondo aplicar pequenos pontos de cor para criar a sensação de cores sólidas (LITTLE, 2010, p. 86-87).

O raciocínio "do ponto à cor difusa" pode ser perfeitamente utilizado para cartografias mais estilizadas. Qualquer formato pode ser utilizado como ponto, mas é comum recorrer a formatos mais simples: quadrados, círculos, triângulos. A densidade dos pontos também pode ser usada como variável para determinar frequências em determinadas regiões.

Figura 21: Uma tarde de domingo na ilha de Grande Jatte, de Georges Seurat. Fonte: Wikimedia. Disponível em: <https://pt.wikipedia.org/wiki/Uma_Tarde_de_Domingo_na_Ilha_de_Grande_Jatte#/media/Ficheiro:A_Sunday_on_La_Grande_Jatte,_Georges_Seurat,_1884.jpg>. Acesso em 30 de maio de 2022.



Mapas que não utilizem a densidade de pontos como variável precisam ter cuidado ao determinar a distância entre pontos. Existe um equilíbrio necessário entre usar pontos grandes o suficiente para ficar explícita a intenção de demarcá-los, mas também valer-se de uma quantidade suficiente para que os contornos do território ganhem a precisão exigida por aquela informação.

O mapa elaborado para este trabalho teve como principal referência o mapa elaborado por David Burney para um projeto envolvendo habitações de emergência (PLUNZ; SOTTO, 2008, p. 65-76). Nele, o autor utiliza um mapa-múndi em projeção de Mercator em vários pequenos quadrados, sendo perfeitamente reconhecível (figura 22). Os continentes aparecem em vermelho, enquanto as áreas suscetíveis a inundações aparecem em cinza. Os círculos vermelhos representam os centros de distribuição da proposta e as linhas pretas conectam a rede.

O tamanho dos pontos impede um entendimento preciso da dimensão dessas áreas mas, no contexto, esse parece justamente a qualidade do mapa: o caráter ensaístico da proposta exige uma compreensão genérica. Um mapa mais detalhado poderia tirar o foco da rede e trazer questionamentos sobre a localização exata dos centros de distribuição, o que não era a intenção do autor.

Apesar de o infográfico proposto neste TFG utilizar os pontos, principalmente, como recurso estético, não é possível abordar esse tema sem citar os usos efetivos da fragmentação do território em grânulos de igual tamanho. Iablonski (2022) cita a Teoria dos Lugares Centrais, de Christaller, para justificar a divisão do espaço em hexágonos. “O formato hexagonal (...) permite que os triângulos formados pelos vértices do lugar central se conectem, e permite que cada célula toque (...) outras seis células” (IABLONOVSKI, 2022). Essa divisão é importante para criar padrões territoriais de cálculo que sejam homogêneos.

Um exemplo da divisão em “pixels hexagonais” é a utilização pelo pesquisador Rafael Pereira e a equipe do Ipea no trabalho para determinar o acesso a oportunidades no Brasil. Os mapas dividem a extensão de cidades brasileiras em hexágonos iguais e índices de acesso a transportes, emprego e saúde são calculados com base nessa divisão (PEREIRA et al, 2020). O resultado são cartografias de fácil leitura e com potencial uso para elaboração de políticas públicas direcionadas (figura 23).

O Sul sendo demarcado no mapa ao invés do Norte tem como referência mapas do Labcidade sobre Covid-19 (MARINO et al, 2021). Entendeu-se que esse recurso seria coerente com a temática do mapa, e foi adotado ao longo de todo o trabalho.

Figura 23: Mapa de Fortaleza utilizado como exemplo no portal Acesso a Oportunidades. Fonte: Ipea. Disponível em: <<https://www.ipea.gov.br/acessooportunidades/en/>>. Acesso em 26 de junho de 2022.

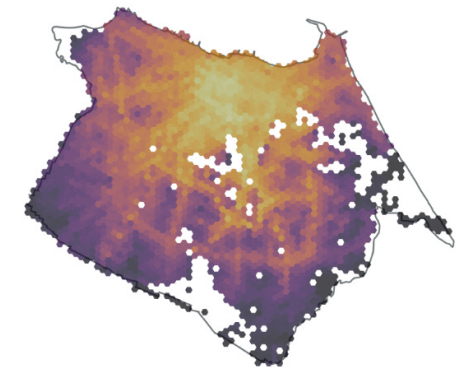
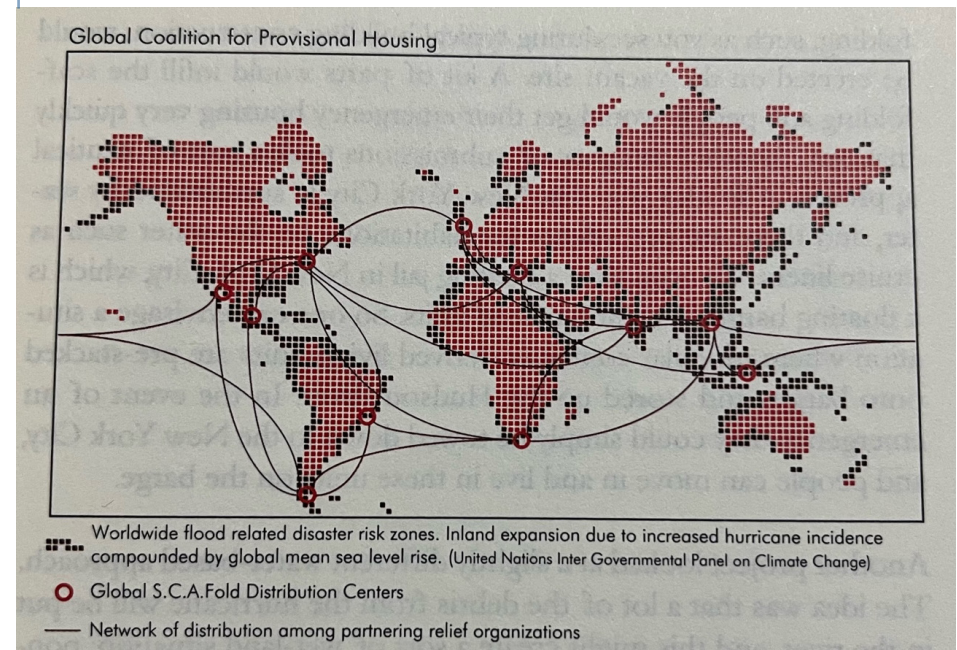
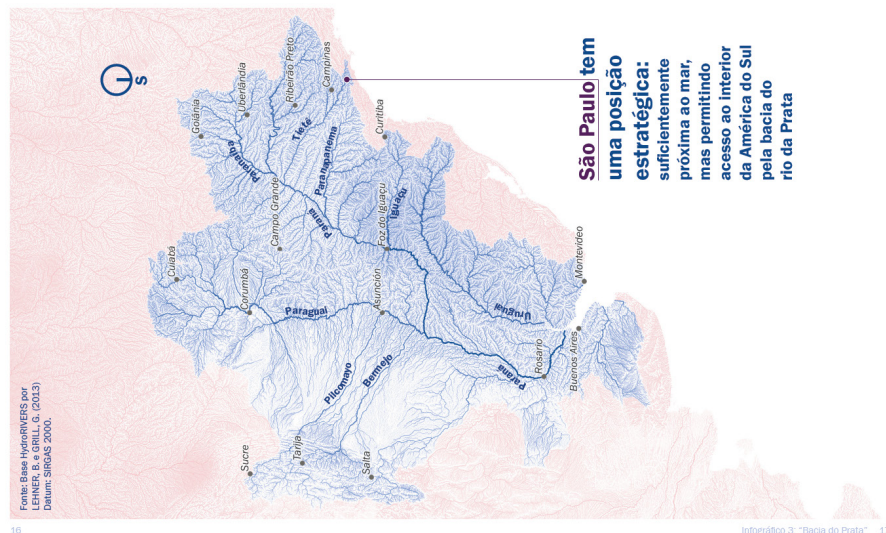


Figura 22: Mapa proposto por Burney. Fonte: PLUNZ; SOTTO, 2008, p. 65-70.



Infográfico 3: Bacia do Prata



Descrição: Mapa mostrando rios na América do Sul, destacando em azul aqueles pertencentes à bacia do Prata. Estão nomeados os rios mais importantes da bacia, assim como algumas cidades.

Objetivo: Contextualizar São Paulo dentro de um contexto hídrico maior. Justificar a própria importância geográfica de São Paulo enquanto ponto de início de expansão para o interior do continente.

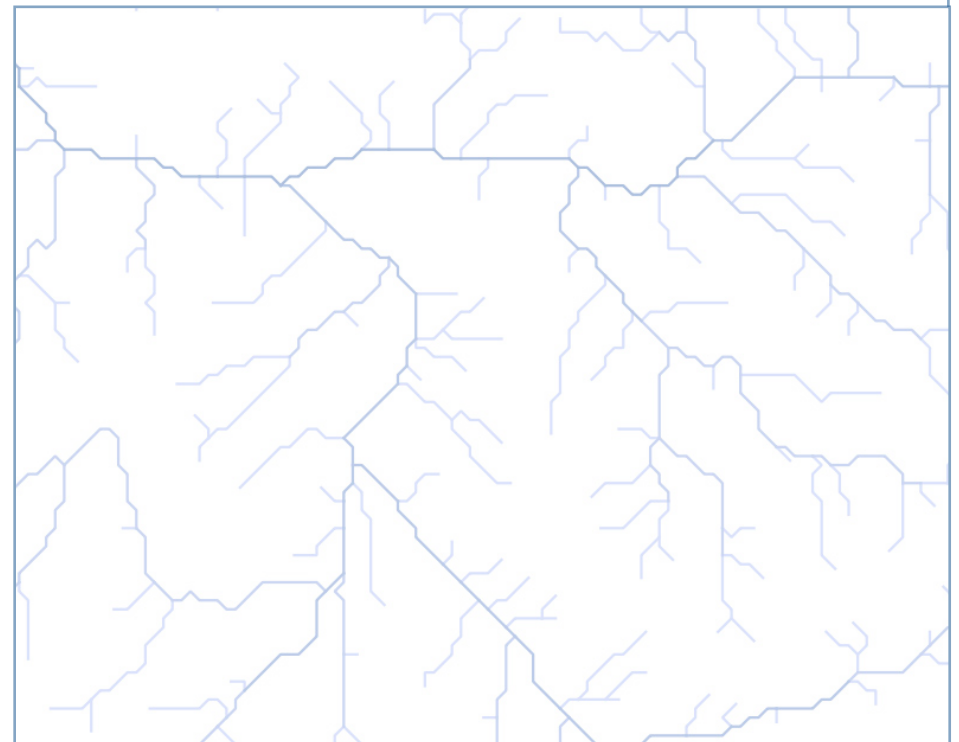
Fonte: HydroSHEDS (bacias) e HydroRIVERS (rios). Natural Earth como base para lista de cidades.

Metodologia: Mapa elaborado em QGIS, incluindo rótulos das cidades. Textos dos rios inseridos diretamente no Indesign.

Discussão: O projeto HydroSHEDS consiste em uma série de bases de dados globais de corpos hídricos. O trabalho extenso é possível por conta de análises e modelagem a partir de imagens de satélite (LEHNER; GRILL, 2013). Uma dessas bases, a HydroRIVERS, consiste nas linhas vetorizadas de todos os rios do mundo com uma bacia de pelo menos 10 km², ou um fluxo médio de 0,1 m³/s, ou ambos.

Como é feita a partir de fotografias de satélite, ela guarda resquícios característicos do trabalho com imagens raster, como a aparência geométrica dos rios ao analisá-los com zoom maior. No caso da região de São Paulo, apesar de rios como o Pinheiros e o Tietê serem facilmente identificáveis, é notável que eles só são divididos em ângulos de 90 ou 45 graus (figura 24). Assim, a base não substitui dados nacionais e regionais de corpos hídricos, mas facilita a compreensão deles em escalas maiores e em regiões com menos produção de dados.

Figura 24: Zoom na região do município de São Paulo. Fonte: O autor, com base em HydroRIVERS (2019).



A abordagem minimalista de um mapa hídrico é inspirada pelo trabalho do T-Map Studio, que faz mapas de rios de diversos países, deixando o traçado das fronteiras subentendido (figura 25). Quanto mais irrigada é a região, mais fácil fica o reconhecimento dos limites; por isso, o sudeste da América do Sul é uma região que tende a parecer um corpo irrigado de veias ou as nervuras de uma folha.

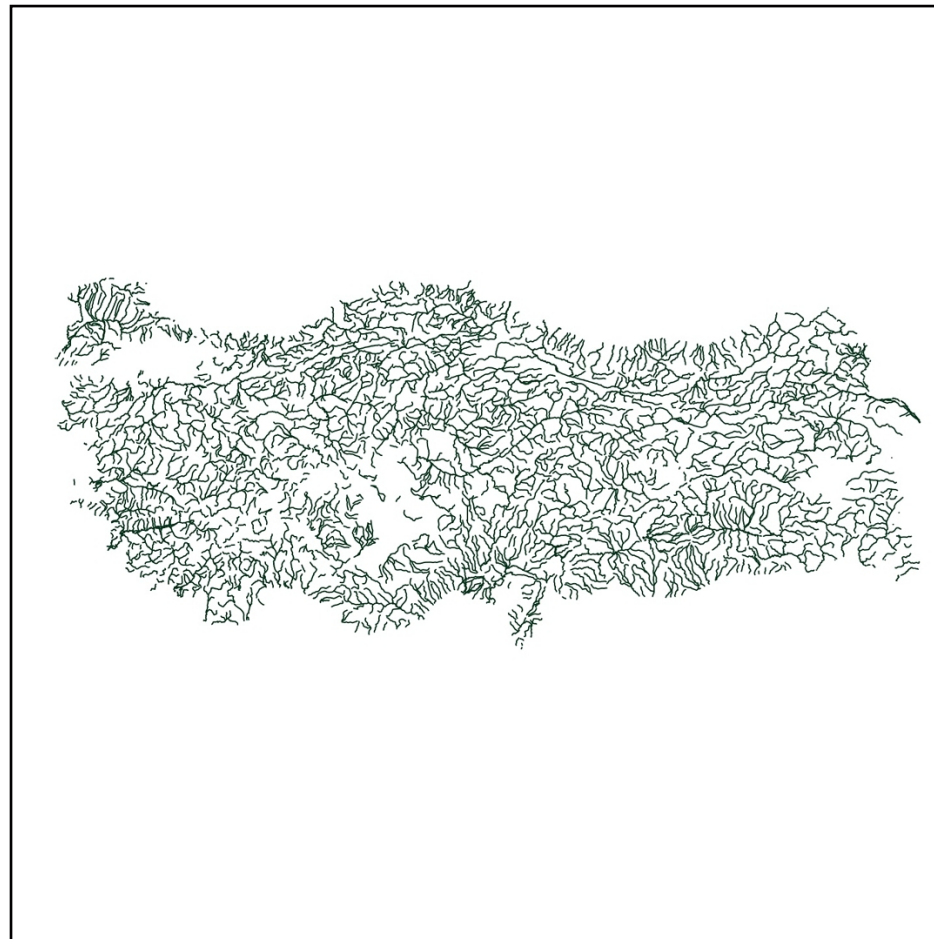
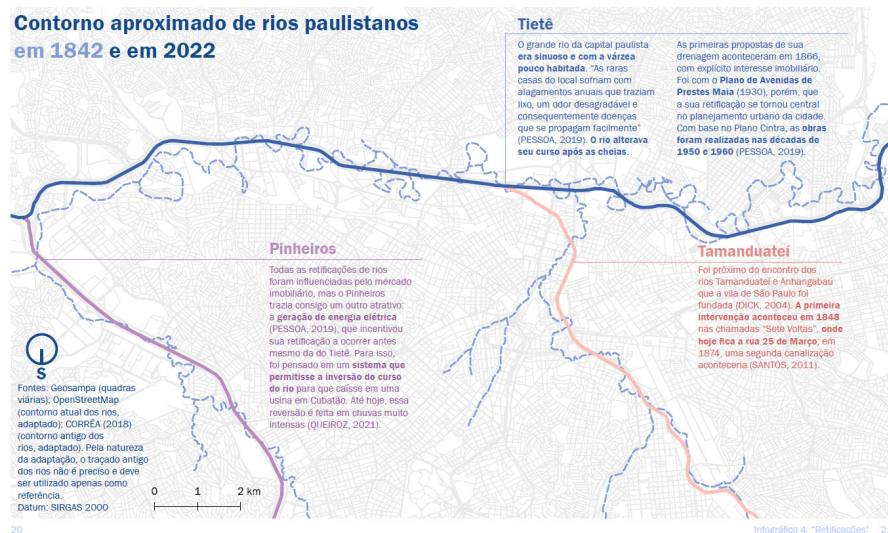


Figura 25: Mapa dos rios da Turquia. Fonte: T-Map Studio, via Instagram. Disponível em: <https://www.instagram.com/p/CN72zo_McKt/>. Acesso em 26 de junho de 2022.

Infográfico 4: Retificações



Descrição: Mapa aproximado de São Paulo mostrando os contornos atuais dos rios Pinheiros, Tamanduateí e Tietê em linha cheia e os contornos em 1842 em linha tracejada. Ao fundo, um mapa com contornos das ruas da cidade.

Objetivo: Mostrar em mapa a diferença dos rios de São Paulo antes e depois da retificação. Dimensionar a grande diferença entre os dois cenários.

Fonte: Os contornos em 1842 têm como referência Correa (2018). O traçado atual dos três rios foi de elaboração própria com base em imagens do Google Maps. O traçado das ruas é do OpenStreetMap, via Geofabrik.

Metodologia: No QGIS, uma imagem do livro "São Paulo: uma biografia gráfica" (CORREA, 2018) foi inserida e georreferenciada. A partir disso, foi traçada a trajetória dos rios em 1842. Para fazer o mesmo com o contorno atual, a imagem do Google Maps já estava georreferenciada pelo plugin HCMGIS.

Discussão: Apesar de a canalização dos rios ser tema central para muitas das discussões de Arquitetura e Urbanismo em São Paulo, conseguir uma base de dados que mostre isso é uma tarefa olímpica. Geosampa, Cesad, agências de águas, órgãos de patrimônio e de urbanismo, meus próprios colegas de FAU USP: ninguém que eu tenha questionado tem um arquivo georreferenciado com esse traçado.

O caminho natural, então, seria pegar um mapa antigo georreferenciado e trabalhar sobre ele. Na plataforma Geosampa, o mais antigo disponível é o Sara Brasil de 1930, que é ótimo, mas não ideal: o Tamanduateí teve suas primeiras obras de canalização ainda em 1848 (SANTOS, 2011). Mesmo assim, o trabalho se mostraria extenso: cada folha do mapa (figura 26) é um arquivo pesado e precisa ser baixado individualmente. Isso se dá por conta do grau de detalhe de cada uma das folhas: algo extremamente importante para análises aproximadas do território, mas desprezível para um mapa mais amplo, que permita menos precisão.

Figura 26: Parte do mosaico de folhas do mapa Sara Brasil. Fonte: Captura de tela do autor, via Geosampa. Disponível em: <http://geosampa.prefeitura.sp.gov.br/PaginasPublicas/_SBC.aspx>. Acesso em 26 de junho de 2022.



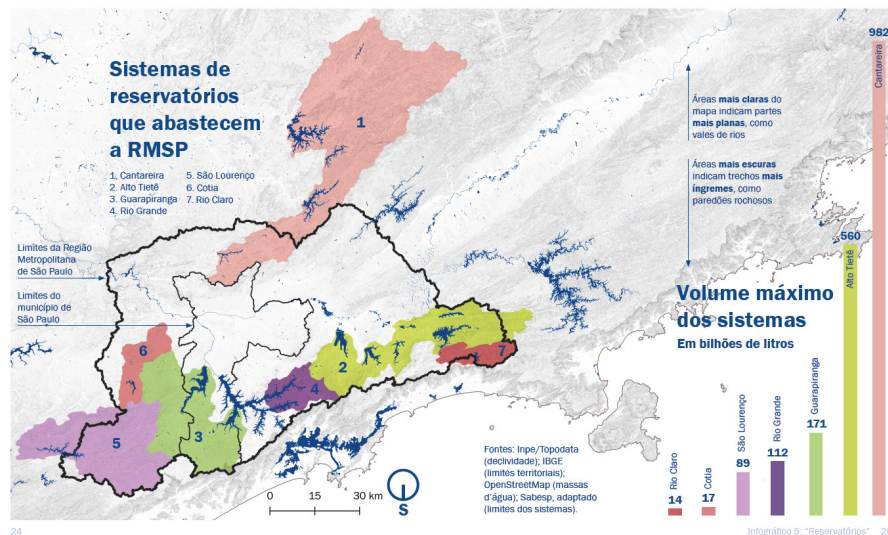
Após pesquisas bibliográficas para outras partes do trabalho, encontrei no livro “São Paulo: uma biografia gráfica”, de Felipe Correa, mapas mostrando a configuração dos principais rios de São Paulo desde 1842 (figura 27), finalizando com um plano para reconectar a cidade com o Tietê. Foi com base nesse mapa que o contorno antigo dos rios foi traçado.

O caminho percorrido, meândrico como o antigo Tietê, para conseguir uma base do contorno antigo dos rios instigou uma curiosidade: eu fui a primeira pessoa a tentar fazer isso na FAU USP? Beira o impossível, considerando a quantidade de discussões envolvendo o tema. A inexistência de uma política de compartilhamento dos dados produzidos nas disciplinas faz com que muito esforço se perca assim que o semestre acaba e se torne um retrabalho para os próximos alunos. Seria útil pensar em estratégias para curadoria, retenção e disponibilização desses arquivos.

Figura 27: Mapa de São Paulo em 1842. Fonte: CORREA, 2018, via Somatic Collaborative. Disponível em: <<https://www.somatic-collaborative.com/research/sao-paulo-a-graphic-biography>>. Acesso em 26 de junho de 2022.



Infográfico 5: Reservatórios



Descrição: Mapa mostrando os sistemas de reservatórios da Região Metropolitana de São Paulo, sobreposto a um mapa de declividade. À direita, um gráfico de barras com o volume máximo dos sistemas.

Objetivo: Contextualizar o estado atual dos sistemas de captação de água em São Paulo. Na narrativa, tem o papel de “preparar o terreno” para as discussões sobre escassez hídrica.

Fonte: Inpe/Topodata para o mapa de declividade; limites territoriais a partir da malha municipal do IBGE; OpenStreetMap para as massas d'água; Sabesp para os limites dos sistemas e para os dados sobre armazenamento máximo de cada um.

Metodologia: Mapa elaborado em QGIS. Gráfico de barras elaborado em R usando o pacote ggplot2. Finalização no Indesign.

Discussão: A primeira versão desse mapa tinha um fundo colorido e o nome de cada um dos reservatórios, mas não o nome dos sistemas (figura 28). A partir de uma conversa com o Prof. Leandro Velloso, concluiu-se que era importante acrescentar os volumes de água dos reservatórios: uma represa pode parecer grande num mapa por sua extensão, mas pode contribuir menos para o abastecimento da metrópole por conta de seu volume.

Os dados dos volumes levaram à constatação de que não valia a pena separar a informação quanto ao nível de água de cada represa. Para o consumidor, afinal de contas, o maior interesse é o volume total de um sistema, o indicador que determina se falta ou não água; também é no nível de sistema que apreendemos boa parte das notícias, como ao ouvir que “falta água na Cantareira”.

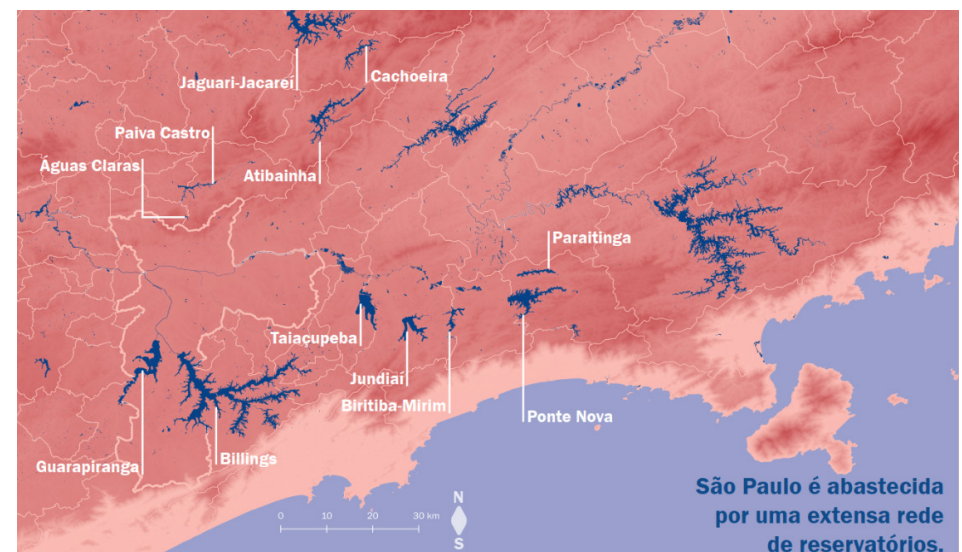


Figura 28: Primeira versão do mapa.
Fonte: o autor.

Os dados dos volumes levaram à constatação de que não valia a pena separar a informação quanto ao nível de água de cada represa. Para o consumidor, afinal de contas, o maior interesse é o volume total de um sistema, o indicador que determina se falta ou não água; também é no nível de sistema que apreendemos boa parte das notícias, como ao ouvir que “falta água na Cantareira”.

Representar os sistemas também significava adotar outra representação de mapa de fundo. O mapa de declividade em tons de cinza e branco permitiu destacá-los e também ajudou a deixar o infográfico com aparência similar aos outros. Tendo sido esse um dos primeiros infográficos elaborados para o trabalho, o processo contribuiu para determinar um elemento de identidade visual comum: o fundo majoritariamente branco.

O fundo branco tem duas funções pragmáticas. A primeira é estética: é muito mais fácil dar uma roupagem mais séria sem ter que trabalhar com cores, além de estar alinhado com a parte de texto corrido do trabalho, que fica em fundo branco. A segunda é técnica: é comum que grandes trechos coloridos fiquem com falhas na impressão.

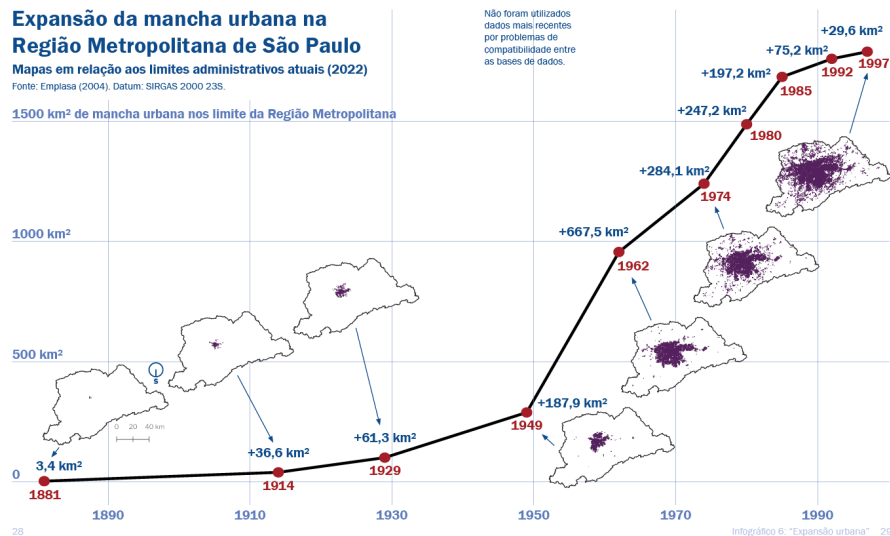
As massas d’água em azul e a textura acinzentada remetem ao mapa presente na reportagem “A Battle Between a Great City and a Great Lake”, do

New York Times. A imagem é usada para contextualizar a importância dos Grandes Lagos e dos canais para o desenvolvimento de Chicago (figura 29).



Figura 29: Mapa da região dos Grandes Lagos, destacando as massas d’água. Fonte: New York Times. Disponível em: <<https://www.nytimes.com/interactive/2021/07/07/climate/chicago-river-lake-michigan.html>>. Acesso em 26 de junho de 2022.

Infográfico 6: Mancha urbana



Descrição: Gráfico de linha mostrando a evolução da mancha urbana de São Paulo entre 1881 e 1997. Mapas acompanham a trajetória da linha.

Objetivo: Enfatizar a rápida velocidade com que a Região Metropolitana de São Paulo se expandiu no século 20, com a consequente impermeabilização de parte expressiva do solo.

Fonte: Emplasa, 2007, por meio do arquiteto Ivan Alves Pereira.

Metodologia: Mapas elaborados em QGIS. Gráfico de linha elaborado em R usando o pacote ggplot2. Finalização no Indesign.

Discussão: Esse simples gráfico de linha foi um dos infográficos mais refeitos de todo o trabalho.

O ponto de partida para a elaboração dele foi uma postagem do arquiteto Ivan Alves Pereira no Twitter (postagem agora apagada) que usava dados sobre a expansão urbana de São Paulo desde a fundação. Mediante solicitação, o autor gene-

rosamente me enviou e informou que se trata de uma base da Emplasa (Empresa Paulista de Planejamento Metropolitano), extinta em 2019. Relatórios disponíveis na Internet comprovam a origem (EMPLASA, 2007).

A proposta inicial de infográfico procurava mesclar as informações da expansão urbana com o contorno antigo dos rios (figura 30). Muita elaboração foi necessária para concluir que seria mais proveitoso considerar dividir essa informação em dois infográficos (além deste, o atual número 4). A decisão de separá-los também tem relação com a compreensão do branco de fundo como parte da identidade visual, como no infográfico 5 (reservatórios).

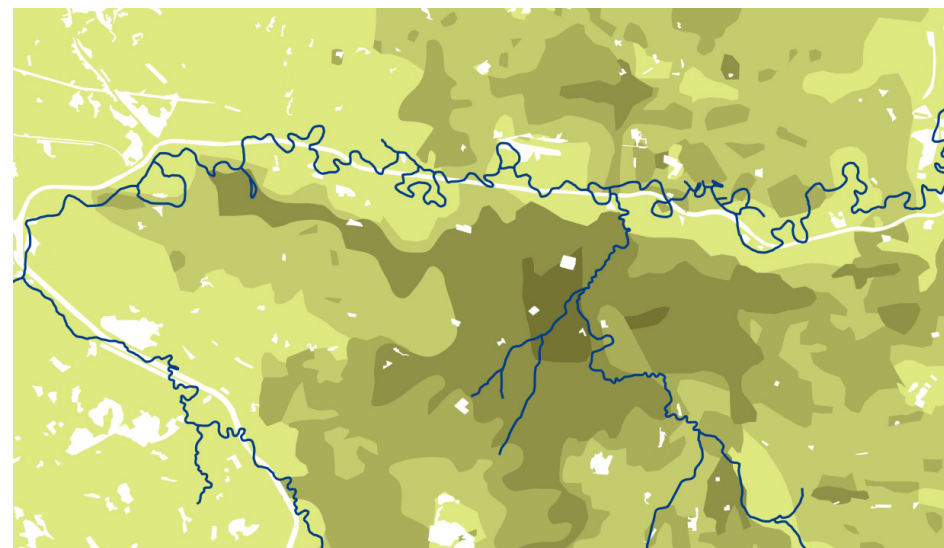


Figura 30: Proposta inicial para mapa de expansão urbana. Fonte: o autor.

A segunda versão do infográfico procurou, com gráficos de barras feitos em R, uma abordagem de visualização matricial (figura 31). O infográfico em formato vertical adotava uma ordem 4x5, em que a primeira coluna era um gráfico com o tempo, as duas intermediárias eram os mapas relativos ao período abordado, e a última consistia em gráfico de barras com o aumento da área.

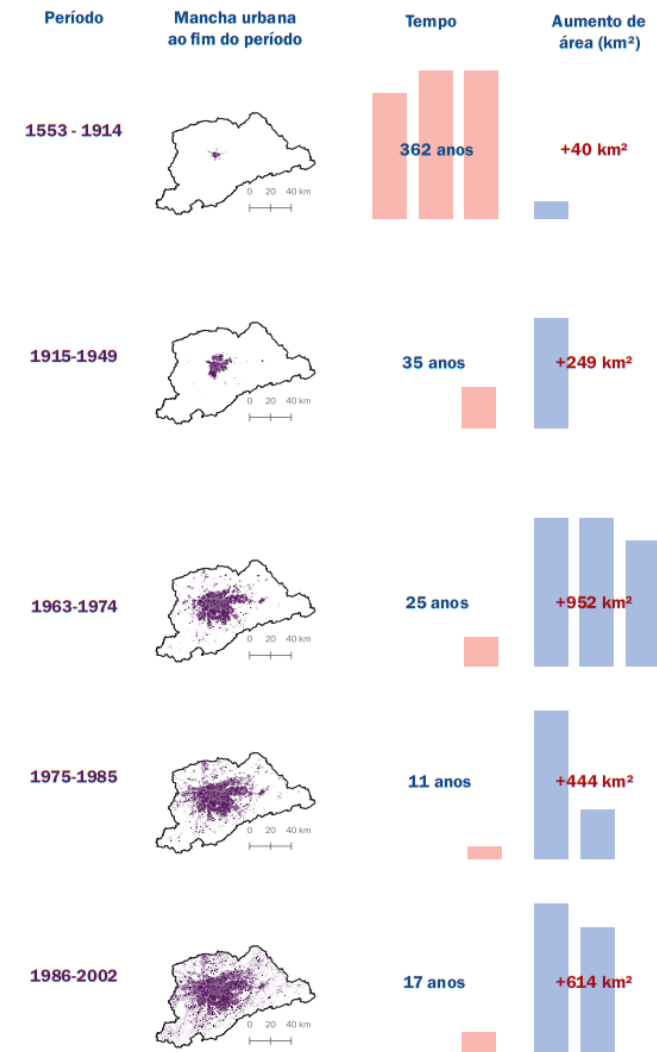
Essa versão acabou ficando confusa para a maioria dos que tiveram contato. Como a variável tempo era muito importante para esse dado, era preferível utilizar uma série temporal. O gráfico de linha, portanto, nasceu disso.

Uma coisa importante a comentar sobre esse dado é que a metodologia utilizada pela Emplasa para registrar o ano de 2002 variou muito em relação aos demais, valendo-se de imagens de satélite de alta resolução (EMPLASA, 2007). O resultado foi uma explosão na área ocupada entre 1997 e 2002 (figura 32) que não encontra embasamento na realidade: pela inclinação da reta, o período de quatro anos teria sido o de urbanização mais acelerada. Por isso, o ano de 2002 não foi incluído na análise, nem outras possíveis análises baseadas em imagens de satélite.

26

Expansão da mancha urbana na Região Metropolitana de São Paulo
Em relação a seus contornos atuais (2022)

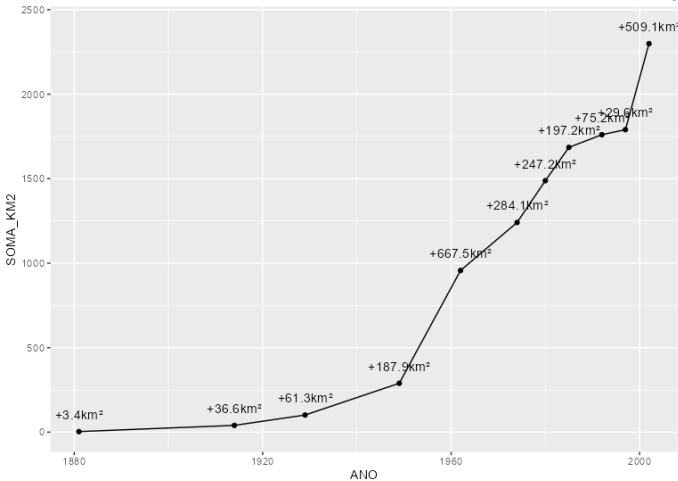
Fonte: Emplasa (2004). Datum: SIRGAS 2000 23S.



Infográfico 6: "Expansão urbana" 27

Figura 31: Segunda versão do infográfico de expansão urbana. Fonte: o autor.

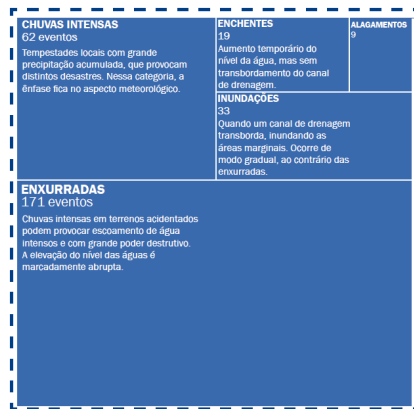
Figura 32 Gráfico "cru" incluindo a variação entre 1997 e 2002 dentro do dataset. Fonte: o autor.



Infográfico 7: Desastres

A maior parte dos desastres no estado de São Paulo são relacionados à água

Registros de desastres no estado de SP entre 2003 e 2016



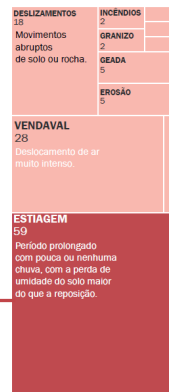
Quando a água ultrapassa a capacidade de escoamento e acumula nas ruas e calçadas (não é desastre natural).

70,2% dos desastres do período eram relativos a grandes volumes de água*,

29,8% não eram, e 14,1% eram relativos justamente à escassez de água

*Deslizamentos podem ocorrer por conta de excesso de água, mas não necessariamente; por isso, foram incluídos como "Outros". Desastres envolvendo gelo também entraram como "Outros".

Fonte: Ministério do Desenvolvimento Regional, Secretaria Nacional de Proteção e Defesa Civil.



Infográfico 7: "Desastres" 33

Descrição: Treemap com os desastres registrados pela Defesa Civil no estado de São Paulo entre 2003 e 2016. À esquerda, em azul, o conjunto de desastres relacionados à água; à direita, em vermelho, os demais.

Objetivo: Destacar o impacto da água nas ocorrências de desastres.

Fonte: S2ID (Sistema Integrado de Informações sobre Desastres) do Ministério do Desenvolvimento Regional.

Metodologia: Gráfico feito em R com os pacotes ggplot2 e treemapify.

Discussão: Os dados de desastres não são simples de se trabalhar. Na base do Ministério do Desenvolvimento Regional, as informações vão só entre os anos de 2003 e 2016. Existem, dentro da mesma plataforma, alternativas para obter dados anteriores, mas de relatórios em formato PDF. Numa pesquisa extensa, que se debruçasse unicamente sobre esses dados, seria factível tabular manualmente essas informações; sendo este um único infográfico num universo maior, o esforço seria muito grande para um resultado que ainda não satisfaria o mais importante: dados atualizados.

Como se trata dos "Reconhecimentos Federais de Situação de Emergência e Estado de Calamidade Pública", a cidade de São Paulo aparece sub-representada. Uma catástrofe na capital paulista precisa ser muito mais intensa do que em outros municípios para estar nessa lista, justamente pela dimensão da metrópole. Assim, é necessário analisar essa base numa esfera maior: optou-se pela estadual, que não representa com precisão o município, mas dá indícios de que a água é um fator importante nos desastres da região.

O gráfico, chamado treemap, é um formato adequado para representar a distribuição dentro de um conjunto de observações. Segundo Friendly (2002), a ideia da representação de dados por um gráfico de área surgiu nos anos 1780 na França. O "Tableau Poleométrique", de Charles de Fourcroy, compara a população de várias cidades europeias da época através de formas retangulares (figura 33). Apesar de não ter o agrupamento "lógico" de um treemap (a soma das partes não gera um todo coeso, só representa muitas cidades desconexas), o precedente visual é claro.

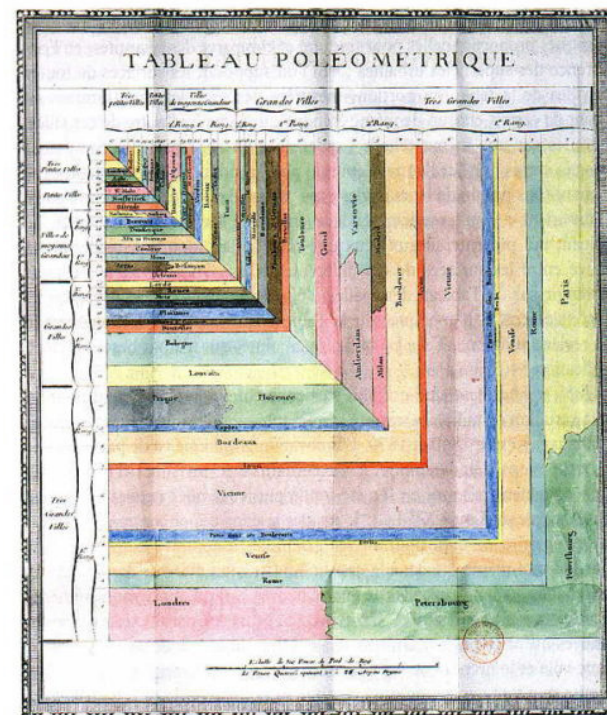
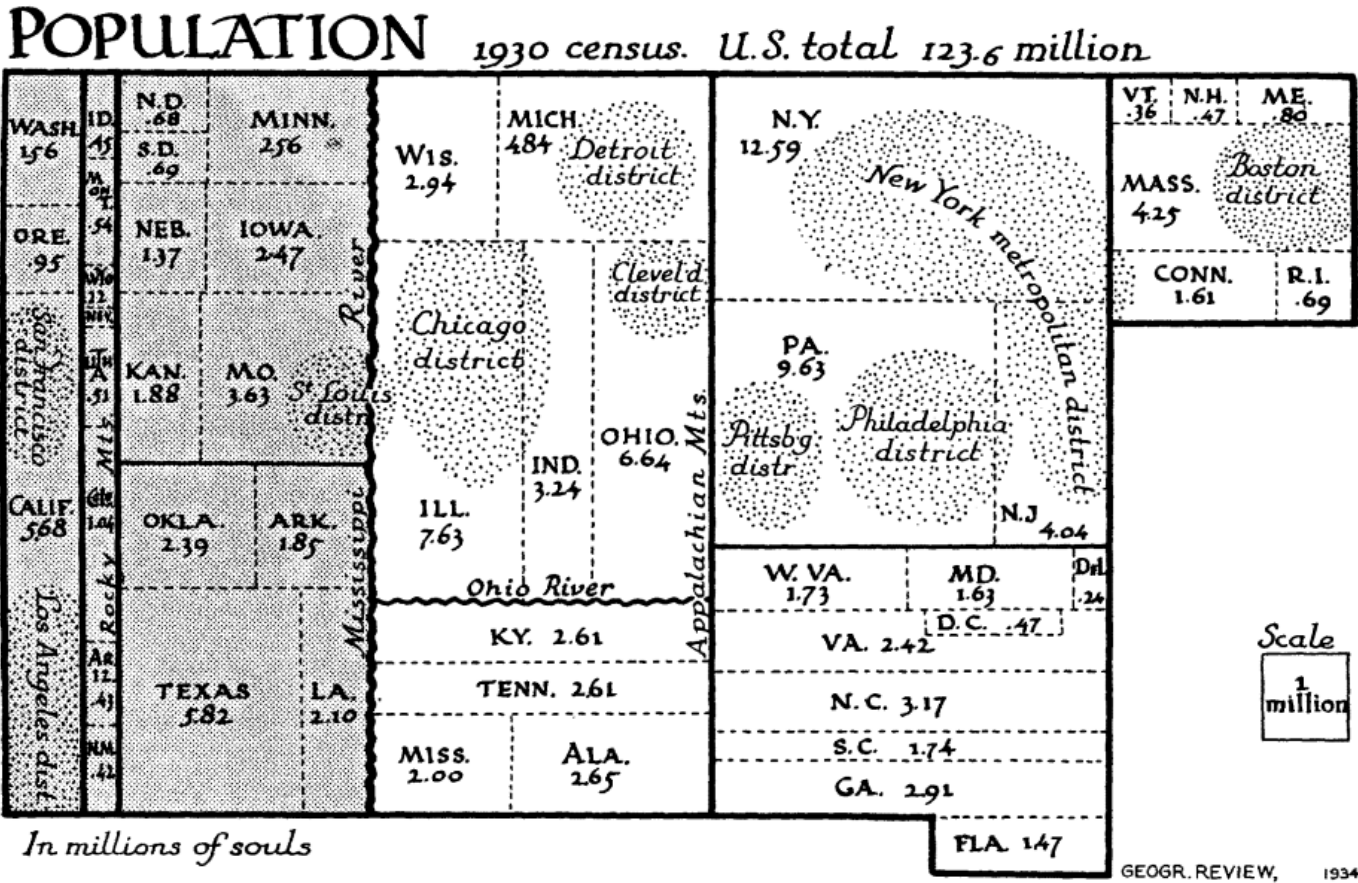


Figura 33: Tableau Poleométrique, de Fourcroy. Fonte: MEKHATRIA, 2015. Disponível em: <<https://www.highcharts.com/blog/tutorials/156-tableau-poleometrique/>>. Acesso em 28 de junho de 2022.

Segundo Scheibel et al (2020), outro precedente dos treemaps são os cartogramas feitos a partir de retângulos, primeiramente usados por Erwin Raisz em 1934. O estatístico seleciona diferentes variáveis e, manualmente, distorce o mapa dos EUA de acordo com elas (o que se dá o nome “cartograma”). É interessante perceber como o autor busca quebrar a abstração da visualização adicionando elementos geográficos, como o rio Mississippi, os montes Apalaches e as próprias dimensões das maiores cidades (figura 34). Nesse gráfico, fica mais claro que as observações formam um todo coeso e que cada uma pertence a um subgrupo.

Os treemaps são uma popular representação em vários campos. Um uso recorrente ocorre com comércio exterior: o OEC (Observatory of Economic Complexity), iniciativa do MIT, utiliza o formato para mostrar as exportações e importações de um país. No exemplo das exportações brasileiras em 2020, em dólares (figura 35), cada produto está agrupando em uma categoria maior representada pela mesma cor, ajudando a entender macrotendências. Minério de ferro e petróleo cru são “produtos minerais”, por isso estão no mesmo marrom alaranjado; soja, milho e café são todos “produtos vegetais”, e assim por diante.

Figura 34: Cartograma retangular com a população dos EUA em 1930. Fonte: RAISZ, 1934. *Geographical Review*, v. 24, n. 2, p. 292, abril de 1934.



[illegible]

Infográfico 8: Aglomerados subnormais

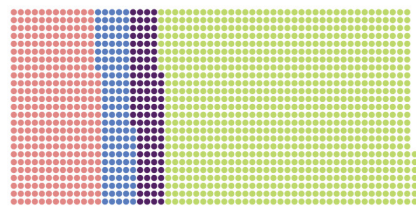
Das 1307 aglomerados subnormais identificados pelo IBGE (2019) no município de São Paulo

317 estão total ou parcialmente em áreas de risco geológico (24,3%)

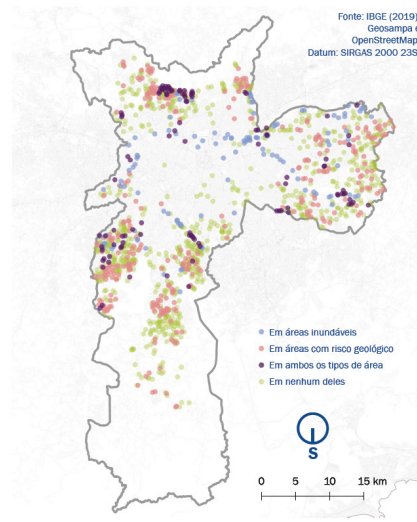
118 estão total ou parcialmente em áreas inundáveis (9,0%)

107 estão total ou parcialmente nos dois tipos de área (8,2%)

765 não estão em nenhuma delas (58,5%), ao menos identificadas até o momento



34



Infográfico 8: "Aglomerados Subnormais" 35

Descrição: Infográfico mostrando os aglomerados subnormais em São Paulo pela localização ou não em área de risco. À esquerda, informação em gráfico tipo waffle. À direita, informação georreferenciada em forma de pontos sobre mapa de ruas.

Objetivo: Contextualizar como a expulsão das populações vulneráveis das áreas centrais também está relacionada à ocupação de áreas de risco.

Fonte: Pesquisa "Aglomerados subnormais", do IBGE (2019). Para as áreas de risco, Geosampa.

Metodologia: Cruzamento entre aglomerados subnormais e áreas de risco utilizando a ferramenta "Seleção por localização" do QGIS. Mapa elaborado em QGIS. Gráfico waffle feito diretamente em Illustrator.

Discussão: Minha primeira tentativa de compreender a territorialização de áreas de risco e populações vulneráveis partiu integralmente do Geosampa. Na plataforma da Prefeitura, existem camadas para risco de inundação, risco geológico, favelas, loteamentos, núcleos e cortiços.

Apesar de os nomes soarem intuitivos, na prática, as categorias são confusas. A figura 36 mostra o entorno da Ponte Estaiada: na primeira imagem, a foto de satélite; na segunda, loteamentos em rosa e favelas em roxo; na última, os aglomerados subnormais do IBGE em preto. A camada de loteamentos inclui territórios de elite no Morumbi e contornos do "loteamento Jardim Edith" distintos de qualquer definição exposta por Lacerda Júnior (2016) sobre o condomínio habitacional de mesmo nome. A mesma figura também revela a desatualização dos dados: a Real Parque ainda aparece na camada de favelas, apesar de a região apresentar atualmente um conjunto habitacional.

Esse descompasso entre a realidade atual e as camadas do Geosampa me levou a buscar uma alternativa. A base de aglomerados subnormais do IBGE contempla vários tipos de habitações em um único termo guarda-chuva, que pode ser sensível para análises mais pontuais, mas que funciona perfeitamente para um exemplo mais amplo. As bases de risco hidro e geológico no Geosampa são de 2021, permitindo uma comparação interessante.

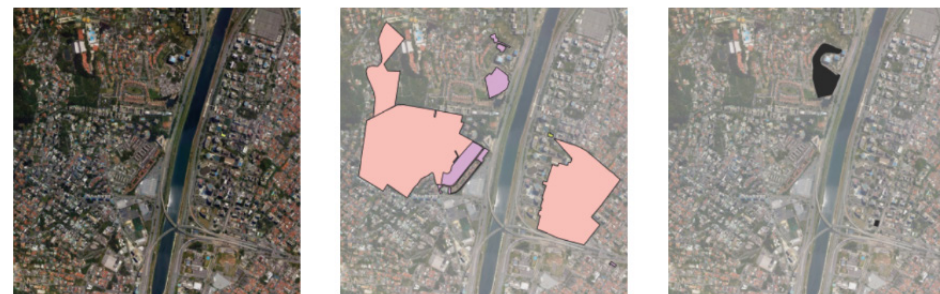


Figura 36: Entorno da Ponte Estaiada.
Fonte: o autor, com base em Google Maps, Geosampa e IBGE (2019).

Em relação à visualização, o mapa de pontos encontra precedente no mapa de John Snow mostrando os focos de cólera em Londres, no século 19 (figura 37). Num período em que ainda se acreditava na teoria miasmática, com as doenças se transmitindo somente pelo ar, o mapeamento permitiu que o médico associasse os casos a uma fonte d'água específica (ROGERS, 2013). O mapa original de Snow não usava pontos para representar os casos: usava pequenos retângulos. A lógica de buscar padrões no território a partir da representação dos eventos em um mapa, porém, tem início nesse trabalho.

Figura 37: Mapa de John Snow.
Fonte: SNOW (2008). Disponível em: <<https://www.thelancet.com/journals/lancet/article/PIIS0140-6736%2808%2960978-2/full-text>>. Acesso em 28 de junho de 2022.



O padrão de pontos se tornou comum por conta da simplicidade da representação. No formato digital, com a facilidade de testar padrões de cor, passou a ser muito utilizado em mapas interativos: quando aproximados, é possível ver detalhes; quando afastados, é possível ver padrões cromáticos. Um famoso exemplo é o Racial Dot Map of America, iniciativa da Universidade de Virgínia que mapeava a população americana a partir de “raça/etnia” (CLARKE, 2022). O site saiu do ar após a divulgação do novo Censo dos EUA, mas existe uma iniciativa brasileira que faz o mesmo com o Censo brasileiro de 2010 (figura 38).

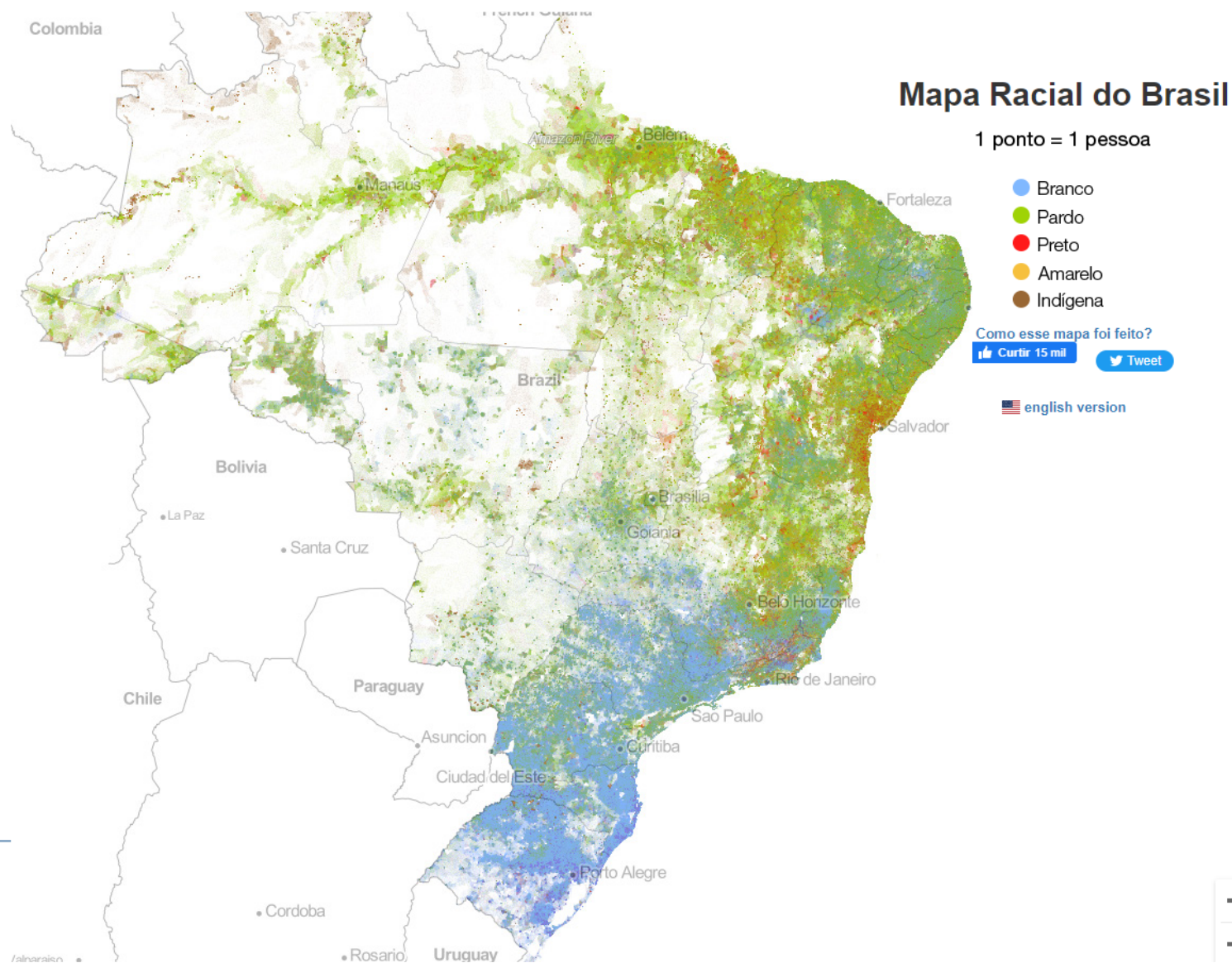


Figura 38: Mapa Racial do Brasil numa escala afastada. Fonte: Pata Data. Disponível em <<http://patadata.org/maparacial/>>. Acesso em 29 de junho de 2022.

Em relação ao gráfico da página esquerda do infográfico, o nome “gráfico waffle” não é um padrão estabelecido, mas o atribuído por Prabhakaran (2017) em sua lista mestra de visualizações do pacote ggplot2, da linguagem R. Ele ganha esse nome por utilizar um grid retangular (figura 39).

O gráfico de Prabhakaran utiliza um grid de 10x10 porque prevê o uso apenas em gráficos de porcentagem: cada quadrado representa 1%. O formato, porém, apresenta uma possibilidade interessante ao dividir o todo em pequenas unidades: é possível utilizá-lo para representar observações individuais (desde que contenham um número não muito grande de observações). Os gráficos utilizados para representar as formações de parlamentos se valem dessa lógica e ainda brincam com o formato curvado, como vários espaços físicos do poder se estruturam (figura 40).

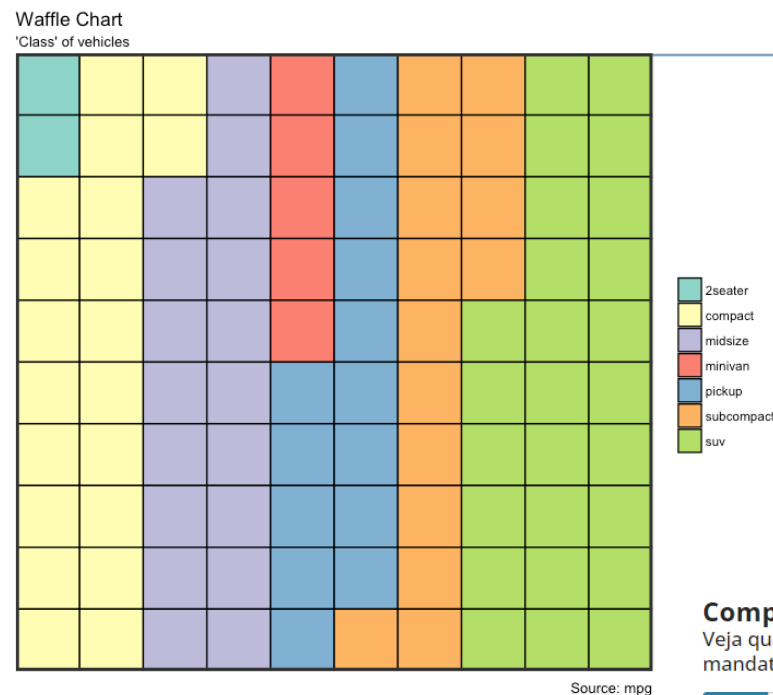


Figura 39: Gráfico waffle. Fonte: PRABHAKARAN, 2017. Disponível em: <<http://r-statistics.co/Top-50-Ggplot2-Visualizations-MasterList-R-Code.html>>. Acesso em 28 de junho de 2022.

Composição da Assembleia Nacional Francesa

Veja quantos deputados têm as diferentes correntes políticas em 2022 e em mandatos anteriores

2022 2017 2012

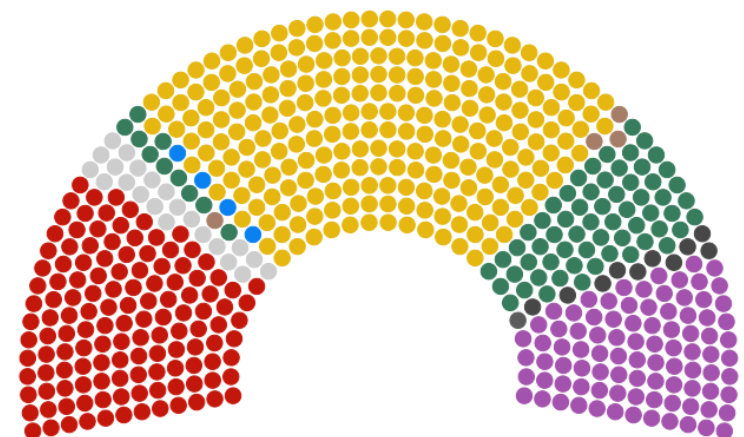


Figura 40: Composição da Assembleia Nacional Francesa por partido. Fonte: G1. Disponível em: <<https://g1.globo.com/mundo/noticia/2022/06/20/veja-grafico-que-mostra-reves-eleitoral-de-corrente-de-emmanuel-macron-na-assembleia-nacional.ghml>>. Acesso em 28 de junho de 2022.

Infográficos 9 e 10: Chuvas

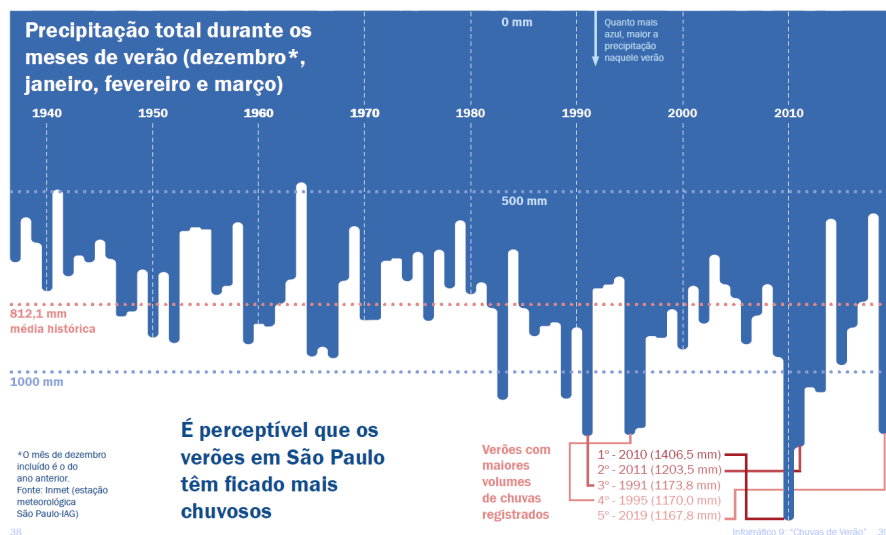


Figura 41: Gráfico de Buache e L'Isle. Fonte: Rendgen, 2020. Disponível em: <<https://sandrendgen.wordpress.com/2020/06/28/history-raise-the-bar-1770s/>>. Acesso em 28 de junho de 2022.

Descrição: Gráfico de barras invertido mostrando a precipitação em São Paulo durante os meses de verão ao longo do tempo. As conexões entre colunas são curvas, lembrando água caindo.

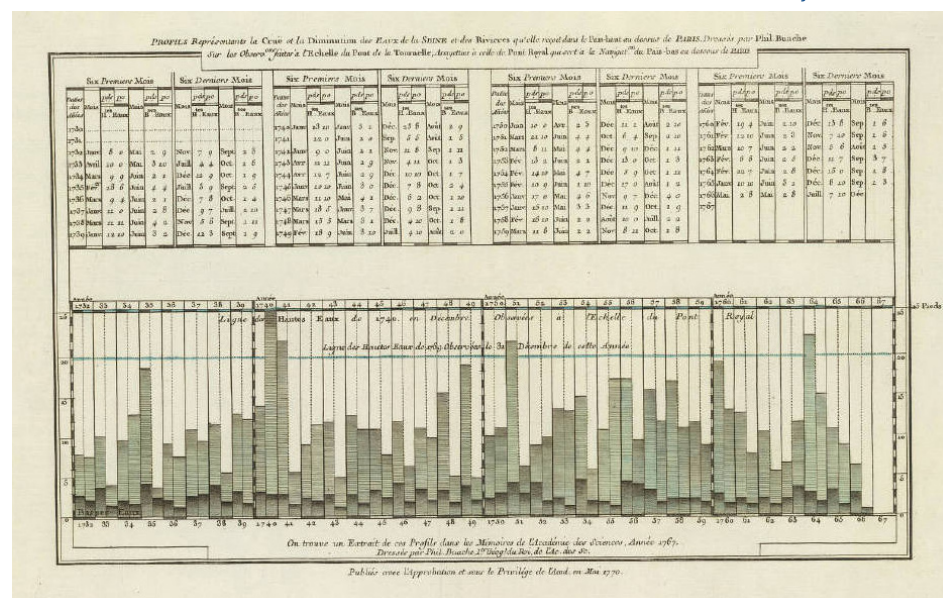
Objetivo: Mostrar a tendência de verões com cada vez mais precipitação em São Paulo.

Fonte: Inmet (Instituto Nacional de Meteorologia). Estação São Paulo-IG.

Metodologia: Gráfico elaborado em R, com o pacote ggplot2. Acabamento em Illustrator.

Discussão: Um gráfico de barras (ou de colunas) é um dos formatos mais populares de gráficos por ser de leitura simples e por permitir a comparação precisa de valores próximos.

O primeiro gráfico de barras que se tem registro, segundo Friendly e Wainer (2022), foi feito por Philippe Buache e Guillaume de L'Isle. Os geógrafos representaram o nível do rio Sena entre 1732 e 1766, num infográfico acompanhado por uma tabela (figura 41).



A criação do gráfico de barras, porém, costuma ser creditada a William Playfair, uma década depois de Buache e L'Isle. Isso acontece porque o gráfico com o nível do Sena é uma transposição direta da realidade: nível é uma variável que representa “altura”, assim como as barras têm uma altura. Playfair abraça a ideia e a extrapola para uma informação mais abstrata: comércio exterior da Escócia. Nas barras mais claras, as importações; nas barras mais escuras, as exportações; todas para 17 países diferentes (figura 42). Uma ideia importante, também, é que os países estão ordenados pela importância de sua participação (FRIENDLY; WAINER, 2022).

O sentido dos gráficos de barras, como indicam uma quantidade, costumam seguir as convenções que entendemos para “maior” (“mais alto”, por isso de baixo para cima) ou “mais pra frente” (no sentido ocidental da leitura, da esquerda para a direita). Seguir convenções, em gráficos, é algo positivo: como nosso cérebro já está acostumado com um determinado jeito de entender o mundo, a compreensão da informação é facilitada (BERINATO, 2019).

Mesmo assim, é justamente a facilidade de entender o gráfico de barras que permite quebrar essas regras. É muito mais fácil convencer o leitor a ter que pensar um pouco mais sobre um gráfico de barras do que sobre outro formato mais complexo.

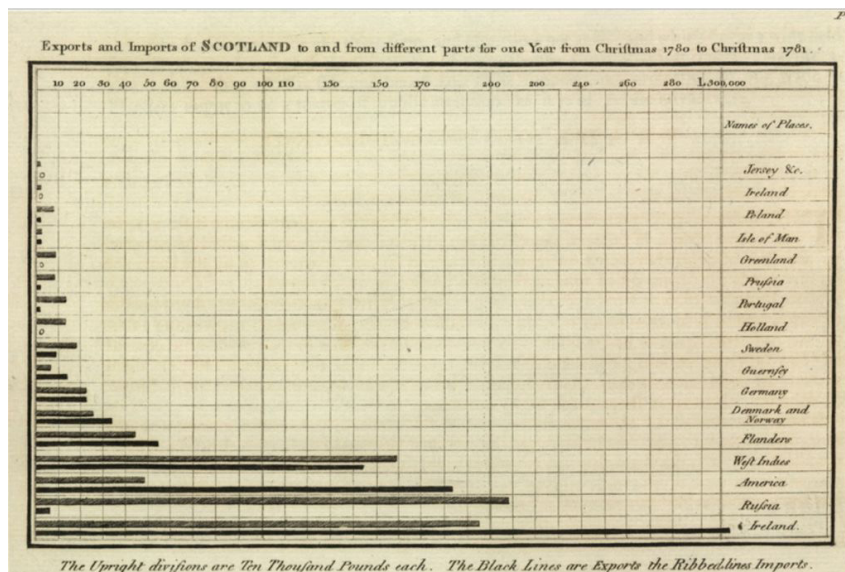


Figura 42: Gráfico de barras de Playfair. Fonte: Rendgen, 2020. Disponível em: <<https://sandrarendgen.wordpress.com/2020/06/28/history-raise-the-bar-1770s/>>. Acesso em 28 de junho de 2022.

Iraq's bloody toll

Um exemplo premiado é o infográfico “Iraq’s Bloody Toll”, de Simon Scarr para o South China Morning Post, de Hong Kong. Ele é um gráfico de barras simples: cada barra corresponde a um mês de baixas na Guerra do Iraque (figura 43). O uso do vermelho, a falta de espaço entre as barras e a inversão do eixo y, porém, fazem com que a imagem pareça uma mancha de sangue escorrendo, reforçando o caráter dramático da informação (HARFORD, 2021).

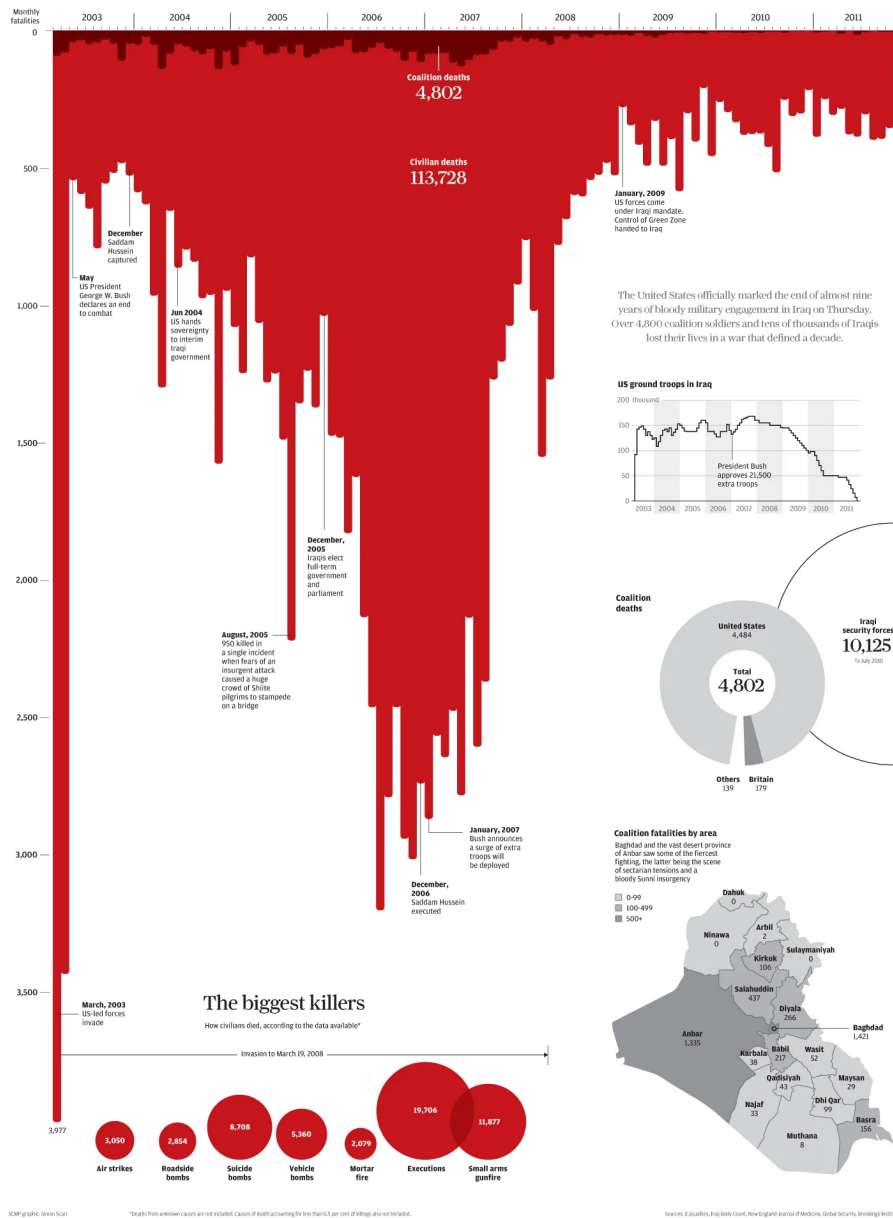
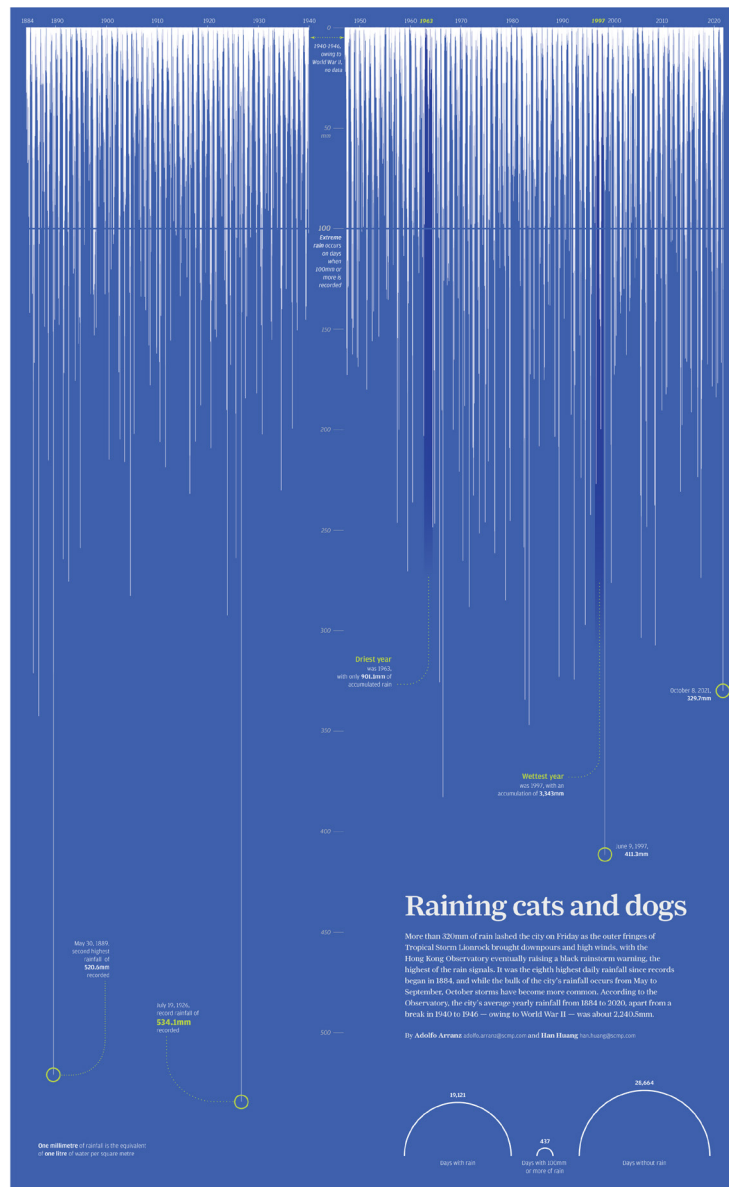


Figura 43: Infográfico de Scarr.
Fonte: South China Morning Post.
Disponível em: <<https://www.scmp.com/infographics/article/1284683/iraqs-bloody-toll>>.
Acesso em 29 de junho de 2022.



O gráfico de Scarr foi a referência mais forte da versão final do infográfico proposto para este trabalho. Num primeiro momento, porém, a inspiração veio de outro material do South China Morning Post. Adolfo Arranz e Han Huang produziram, em 2021, um gráfico de barras invertido mostrando precipitações diárias em Hong Kong (figura 44). O contraste entre branco e azul também fornece a sensação de líquido escorrendo — no caso, a água da chuva.

Os gráficos de barra do infográfico 10, mostrando garoa e tempestade, foram feitos após a leitura do artigo de Marengo et al (2020) analisando os mesmos dados. Um grande número de infográficos com visualizações diferen-

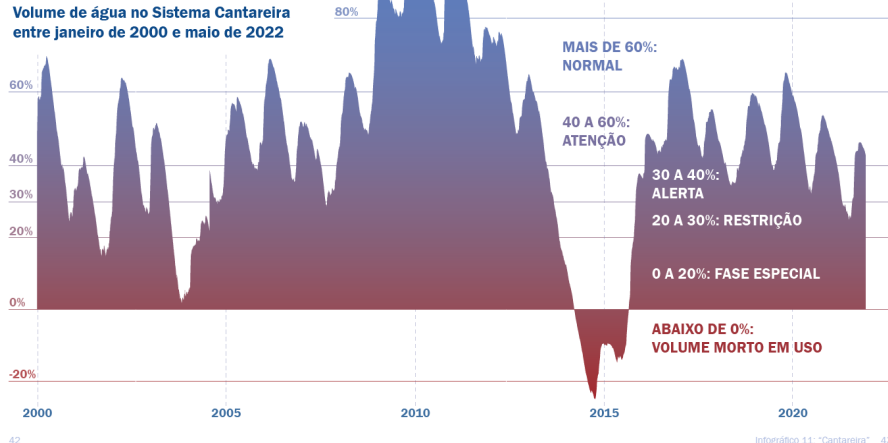
O infográfico da chuva em Hong Kong foca nos dias com maior volume de chuva. Apesar de transparecer mais episódios de chuvas extremas a partir de 1940, não permite compreender tendências de mais longo prazo. No caso da chuva em São Paulo, muitos testes foram feitos com os dados diários. Foi constatado que, ao não mostrar tendências, o gráfico não contribuía para a narrativa geral. Muitos testes foram feitos até o agrupamento da precipitação nos meses de dezembro, janeiro, fevereiro e março se mostrar uma boa alternativa.

Figura 44: Infográfico de Arranz e Huang. Fonte: South China Morning Post. Disponível em: <<https://multimedia.scmp.com/culture/article/SCMP-printed-graphics-memory/lonelyGraphics/202110A294.html>>. Acesso em 29 de junho de 2022.

Infográficos 11: Cantareira

Os reservatórios que abastecem São Paulo não têm conseguido recuperar o volume perdido

Volume de água no Sistema Cantareira entre janeiro de 2000 e maio de 2022



Descrição: Gráfico de linha com área preenchida com degradê mostrando o volume de água disponível no Sistema Cantareira, de 2000 à atualidade.

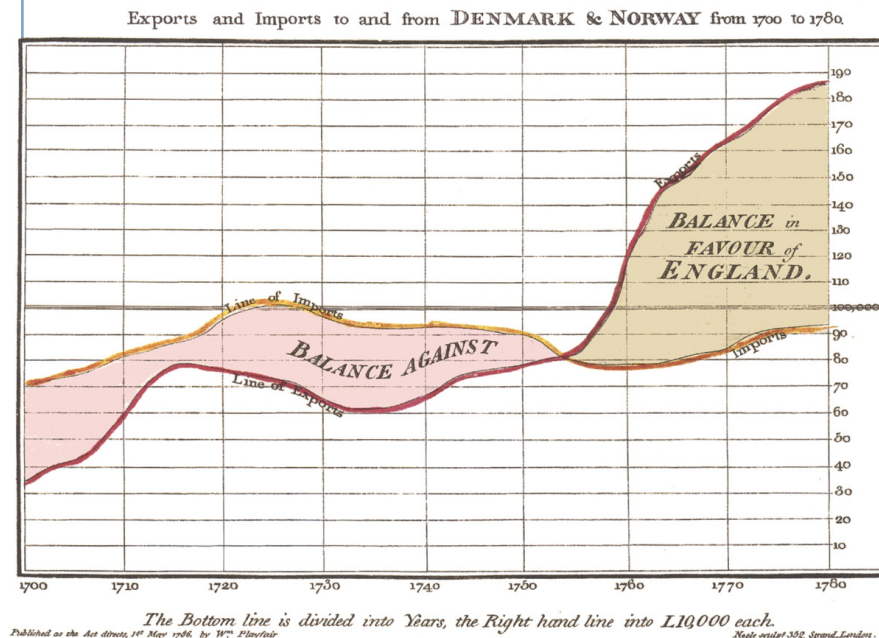
Objetivo: Após a sequência de gráficos mostrando chuvas frequentes, mostrar que também existe o problema da estiagem na mesma cidade.

Fonte: Sabesp, via pacote em R “mananciais” de Beatriz Milz.

Metodologia: Tratamento dos dados e gráfico inicial em R. Degradê adicionado no Illustrator.

Discussão: Gráficos que misturam área e tempo para sinalizar quantidade num intervalo de tempo não são novos. O precedente dessa técnica já aparece nos gráficos de William Playfair (figura 45), feitos no século 18 (BYRON; WATTENBERG, 2008). O gráfico de linhas mostra exportações e importações da Inglaterra com relação à Dinamarca e à Noruega, e o preenchimento da área entre elas corresponde à balança comercial inglesa.

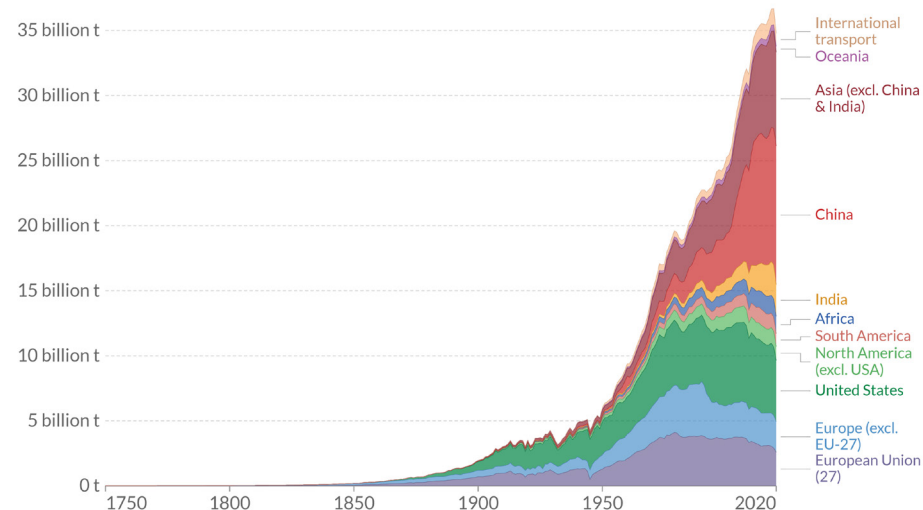
Figura 45: Gráfico de Playfair mostrando a balança comercial inglesa. Fonte: Wikimedia. Disponível em: <https://en.wikipedia.org/wiki/William_Playfair#/media/File:Playfair_TimeSeries-2.png>. Acesso em 26 de junho de 2022.



Um uso bastante popular do preenchimento em gráficos é na comunicação de várias variáveis acumuladas. Na figura 46, são mostradas as emissões de CO₂ por combustíveis fósseis por ano, separadas por grupos de países. O preenchimento com cor tem uma função prática de distinguir cada variável e reforçar que o valor máximo é o acumulado de todas elas; se fossem apenas linhas coloridas, poderia fazer entender ser o valor absoluto de cada uma.

Em relação ao infográfico produzido para este trabalho, a cor busca reforçar os valores que a própria Sabesp estima como sendo de normalidade, atenção e outras fases críticas. As primeiras tentativas foram feitas colorindo os trechos com linhas verticais conforme as fases, mas passou sensação de desordem. Com as cores na horizontal, houve tentativas com cores sólidas (figura 47), mas o resultado também pareceu inadequado. O uso do degradê conseguiu passar a mensagem de “cor representando categoria” de maneira harmônica.

Annual CO₂ emissions from fossil fuels, by world region



Source: Global Carbon Project
OurWorldInData.org/co2-and-other-greenhouse-gas-emissions • CC BY
Note: This measures CO₂ emissions from fossil fuels and cement production only – land use change is not included. 'Statistical differences' (included in the GCP dataset) are not included here.

Figura 46: Gráfico de emissões de CO₂ por combustíveis fósseis. Fonte: Our World in Data. Disponível em: <<https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions>>. Acesso em 30 de junho de 2022.

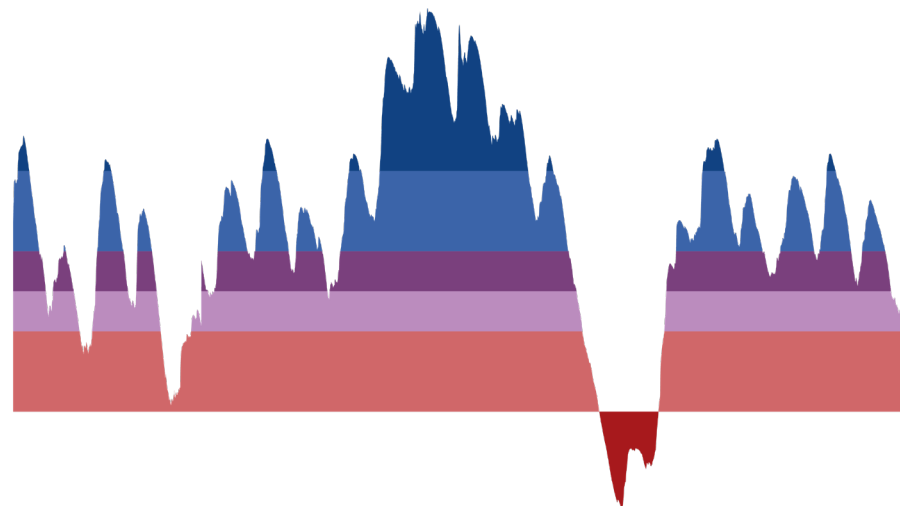


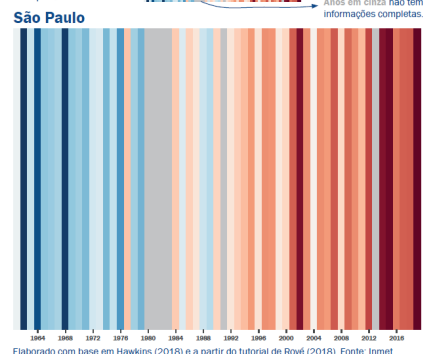
Figura 47: Infográfico dividido em cores sólidas na horizontal. Fonte: o autor.

Infográficos 12: Warming stripes

O aumento da temperatura média anual já é notável em São Paulo atualmente

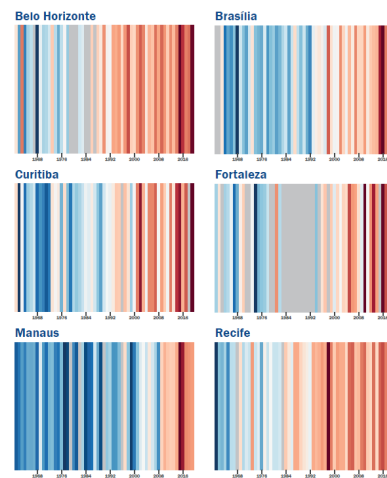
Como ler este gráfico:
Cada coluna representa a temperatura média de cada ano. Quanto mais branco, mais perto da média.

Anos mais azuis foram mais frios que a média.
Anos mais vermelhos foram mais quentes que a média.
Anos em cinza não têm informações completas.



44

Outras cidades enfrentam o mesmo



Infográfico 12: "Warming stripes" 45

Descrição: Gráficos de barras mostrando a temperatura média anual em diversas cidades por meio da cor. A média do conjunto dos anos aparece em branco; quanto mais azul, mais fria em relação à média; quanto mais vermelho, mais quente.

Objetivo: Reforçar que o aquecimento da temperatura média está ocorrendo em São Paulo e em outras cidades.

Fonte: Inmet (estações São Paulo-Mirante, Belo Horizonte, Brasília, Curitiba, Fortaleza, Manaus e Recife).

Metodologia: Com o tutorial de Royé (2018), os gráficos foram gerados em R.

Discussão: O aumento da temperatura global costuma ser representado com gráficos de linha ou de barras. Apesar de serem precisos, não têm apelo visual tão grande.

A ideia de representar essa informação com listras começou no mundo do crochê. Duas cientistas tiveram essa ideia em momentos distintos, independentemente. Joan E. Sheldon, em 2017, fez um cachecol com as temperaturas de 1600 até 2000, utilizando a média de 1600 a 1900 como ponto de referência (figura 48). Já

em 2017, Ellie Highwood fez um cobertor de bebê com as temperaturas de 1916 a 2016 (figura 49) (MUNSELL, 2022). A ideia de utilizar dados climáticos era tendência entre as pessoas que tinham crochê como hobby: já eram populares artefatos representando a temperatura em um lugar específico, assim como representando a variação da cor do céu (HIGHWOOD, 2017).



Figura 48: Cachecol original de Sheldon. Fonte: Sheldon, 2017. Disponível em: <<http://sheldonfiberdesigns.net/the-globally-warm-scarf/>>. Acesso em 30 de junho de 2022.

Figura 49: Cobertor original de Highwood. Fonte: Highwood, 2017. Disponível em: <<https://elliehighwood.com/2017/06/12/climate-changecrochet-the-global-warming-blanket/>>. Acesso em 30 de junho de 2022.



Na postagem original em seu blog, Highwood inclui uma versão digital das listras do seu cobertor (HIGHWOOD, 2017). Na imagem, feita pelo cientista climático Ed Hawkins, é proposta uma escala que fosse acessível a pessoas com daltonismo (figura 50).

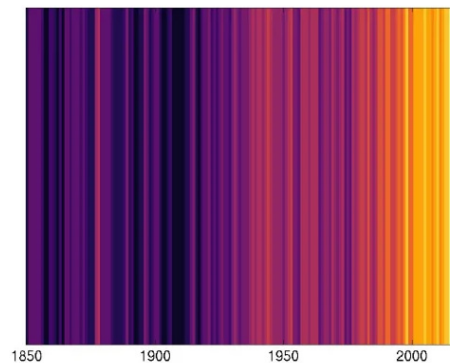


Figura 50: Primeiras listras climáticas de Ed Hawkins. Fonte: Highwood, 2017. Disponível em: <<https://elliehighwood.com/2017/06/12/climatechangecrochet-the-global-warming-blanket/>>. Acesso em 30 de junho de 2022.

Em 2018, Ed Hawkins foi chamado para falar em um festival de literatura sobre o aquecimento do planeta. Utilizou o mesmo gráfico feito para Highwood, modificando as cores para azul e vermelho, associadas a frio e calor (MUNSELL, 2022). O gráfico teve uma boa recepção do público, principalmente após ser divulgado em redes sociais. A figura 51 foi retirada do site Show Your Stripes, que permite explorar as “listras climáticas” de diversos países.

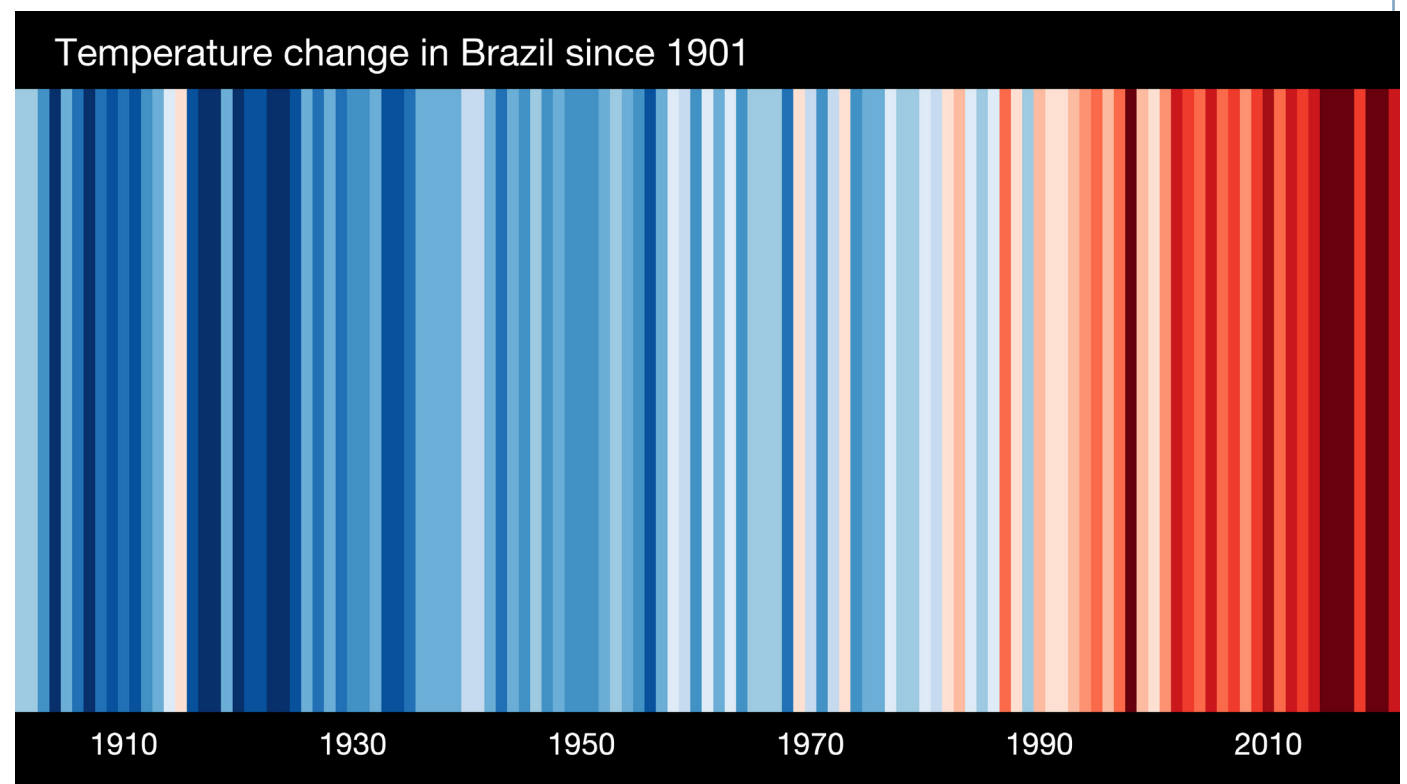


Figura 51: Listras climáticas para o Brasil. Fonte: Show Your Stripes. Disponível em: <<https://showyourstripes.info/l/southamerica/brazil/all>>. Acesso em 30 de junho de 2022.

A repercussão do formato de gráfico foi grande. As listras foram reproduzidas na moda (figura 52), em arquitetura, em obras-de-arte; o primeiro livro de Greta Thunberg, ativista pelo clima, tem o gráfico como capa. É um exemplo interessante de como as pessoas podem se conectar sensorialmente com informações técnicas através de uma boa visualização de dados.

É preciso comentar que esse era um dos gráficos que eu mais queria fazer. A primeira vez que me deparei com o tutorial de Royé (2018) foi no final de 2021, e fiquei assustado com as linhas de código que eu não conseguia entender direito. Quando reli o tutorial em junho de 2022, foi muito menos desafiador do que parecia. Esse teste acabou sendo um atestado de que estou entendendo mais de programação do que há seis meses.

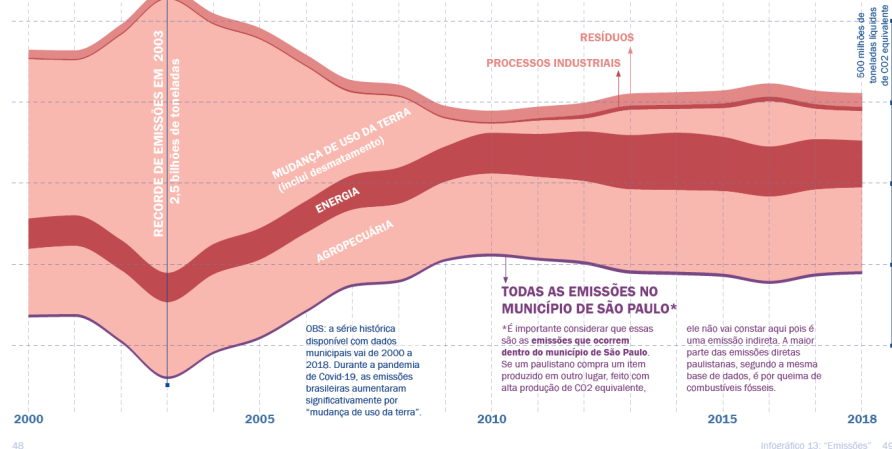
Figura 52: Modelo na London Fashion Week usando roupa com listras climáticas. Fonte: Hawkins, 2022. Disponível em: <https://twitter.com/ed_hawkins/status/1539182461455843328>. Acesso em 30 de junho de 2022.



Infográficos 13: Emissões

Emissões de gases estufa entre 2000 e 2018 no Brasil

Fonte: SEEG, via basedados.gov.br



Descrição: Gráfico do tipo “streamgraph” agrupando as emissões líquidas de CO₂ equivalente no Brasil, por categoria. As emissões do município de São Paulo estão destacadas em relação ao restante.

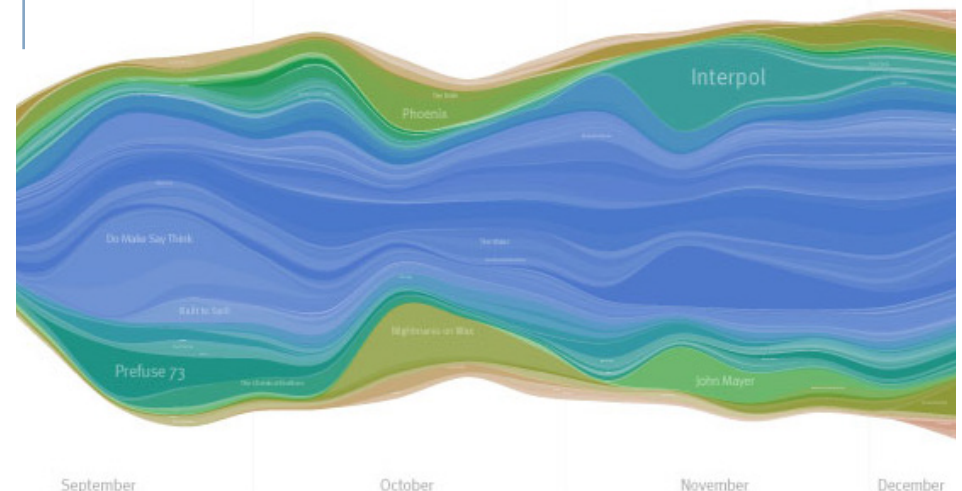
Objetivo: Representar a participação de São Paulo nas emissões líquidas nacionais de CO₂ equivalente. Reforçar o argumento de que as emissões de CO₂ são um problema em escala muito maior do que a municipal.

Fonte: Buscou-se, inicialmente, trabalhar diretamente com os dados do Sistema de Estimativa de Emissões e Remoções de Gases de Efeito Estufa (SEEG). A complexidade da base levou à alternativa dos dados já tratados na iniciativa Base dos Dados.

Metodologia: Por meio da linguagem de programação R, agrupou-se primeiro os dados do município de São Paulo (código IBGE 3550308) por ano. Os dados restantes foram agregados segundo a categoria “Nível 2” da tabela. Foram consideradas as emissões líquidas, ou seja, incluíram-se as absorções de CO₂ equivalente pela vegetação. Para a confecção do gráfico, foi utilizada a plataforma RAWGraphs, com finalização no Illustrator.

Discussão: Como já analisado no infográfico 11, gráficos de linhas com áreas pintadas têm precedentes no século 18. O streamgraph, porém, revoluciona a técnica ao se desprender do eixo “quantidade” iniciando no zero. O primeiro registro é de um trabalho de Lee Byron chamado Listening History (figura 53) que mostra o próprio consumo de música temporalmente com base em dados do last.fm (BYRON; WATERNBERG, 2008). Quanto maior o trecho, mais vezes o artista foi ouvido; a cor corresponde à primeira vez que ouviu, indo do azul (mais antigo) a um tom de areia (mais recente).

Figura 53: Trecho do gráfico original do Listening History.
Fonte: Ben Fry, 2006. Disponível em: <<https://benfry.com/fall06/lee/>>. Acesso em 26 de junho de 2022.



O formato se popularizou após uma matéria do New York Times chamada “The Ebb and Flow of Movies: Box Office Receipts 1986-2007”, que mostrava a arrecadação de filmes nos cinemas dos EUA. Nele, quanto maior a área, maior o lucro (figura 54). As cores são a receita total do filme, mas como representam variáveis discretas (apenas quatro), têm uma leitura muito mais simples do que a do gráfico de música (BYRON; WATERNBERG, 2008). A maioria dos filmes inicia com um formato alongado e vai se achatando, deixando claro que a receita nas semanas iniciais é maior e vai progressivamente diminuindo. A área total mostra quando as pessoas vão mais ao cinema.

O formato apresenta limitações. A ordem dos elementos no gráfico pode causar distorções, principalmente no caso de “saltos” muito repentinos (DI BARTOLOMEO; HU, 2016). A ausência do eixo de referência iniciando no zero impossibilita a leitura dos valores com precisão. Assim, só se deve utilizá-lo quando os valores individuais forem menos importantes que a narrativa geral; o uso de legendas com os valores pode minimizar esse problema (figura 55).

Apesar disso, o gráfico costuma ter um bom retorno de público (BYRON; WATERNBERG, 2008) e também ajuda a dar uma roupagem menos sisuda para um conjunto de dados.

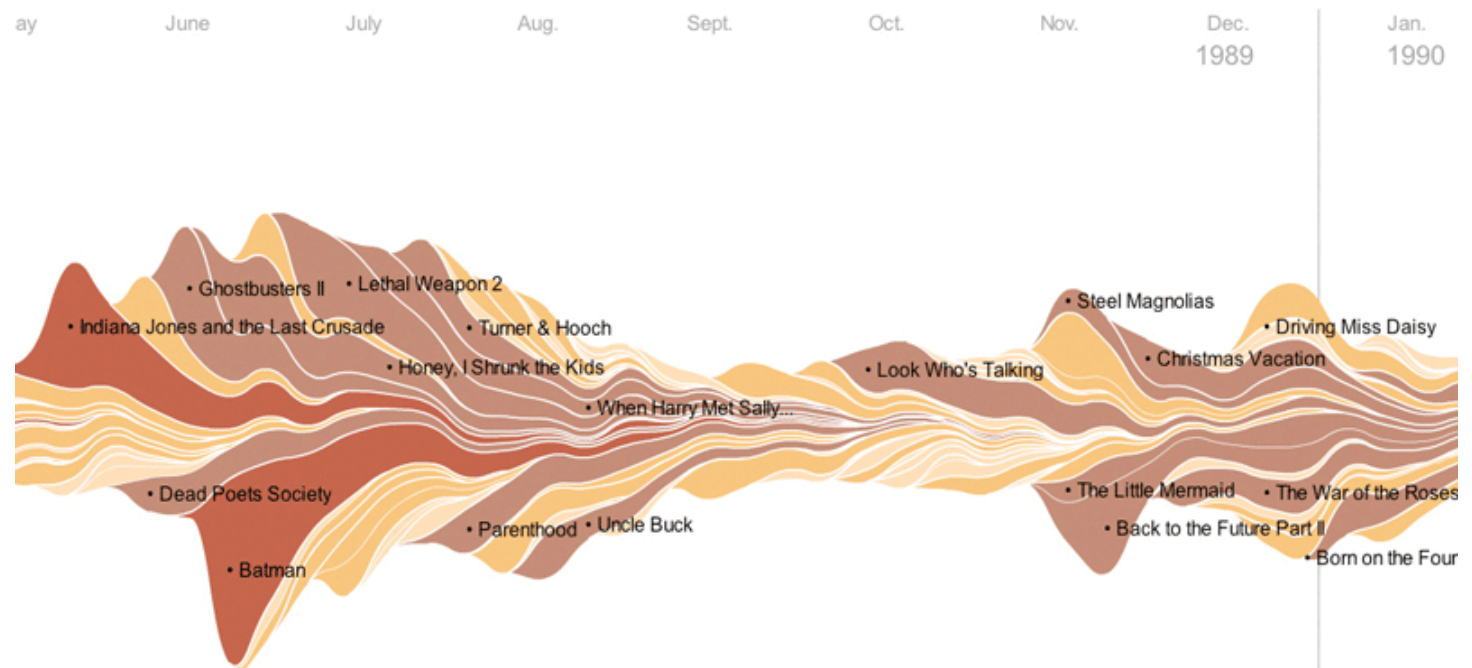


Figura 54: Trecho do gráfico do New York Times mostrando a arrecadação de filmes nos EUA. Fonte: A Canary in the Data Mine. Disponível em: <<https://www.katiepeek.com/data-minecanary/2010/01/movie-box-office-figures-from-1986-to-2008-from-the-new-york-times.html>>. Acesso em 26 de junho de 2022.

Número de novos apartamentos por bairro, segundo o ano de construção

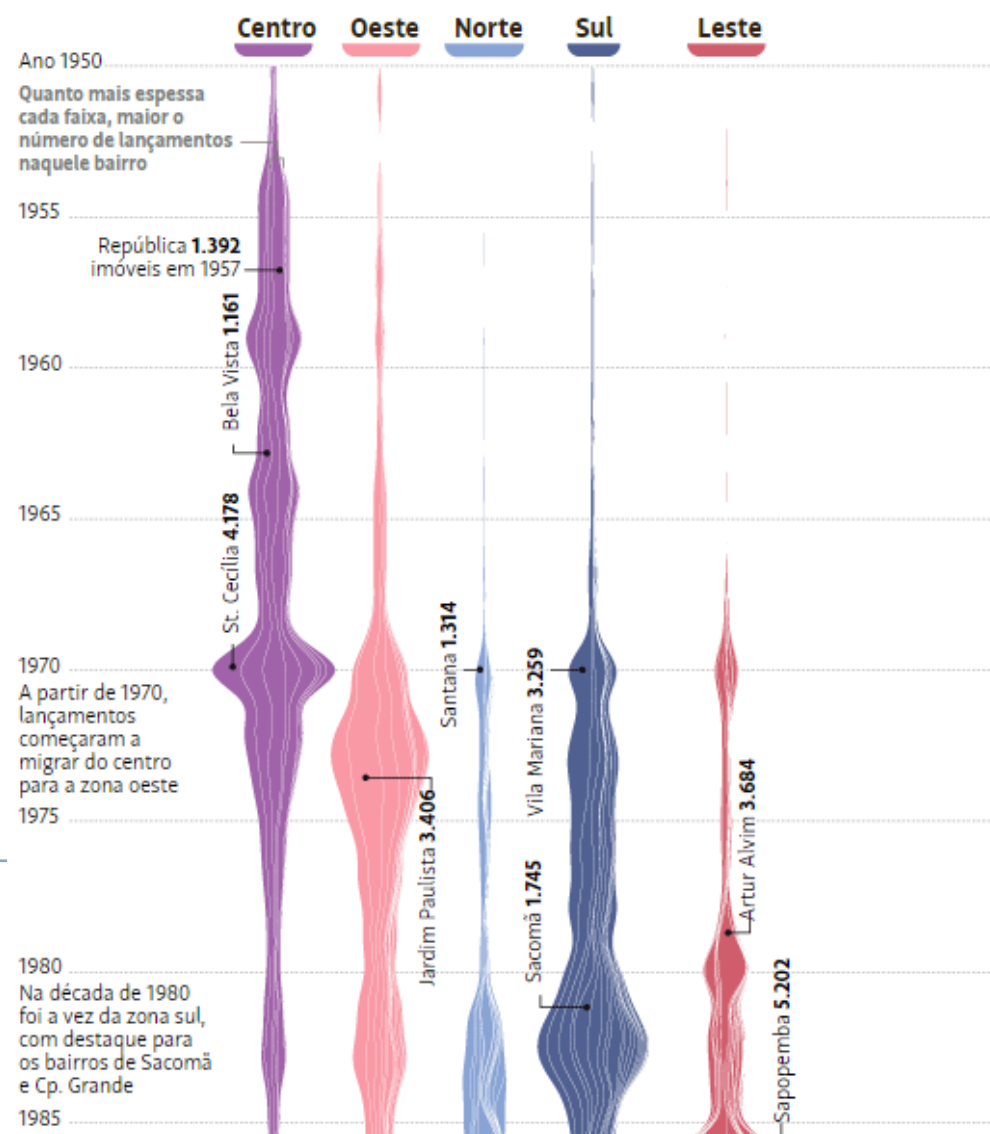
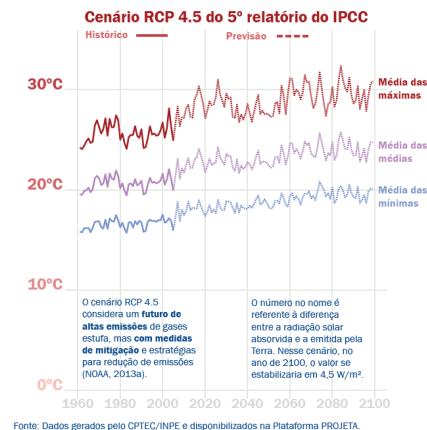


Figura 55: Trecho de streamgraph mostrando o número de apartamentos novos por região de São Paulo, por ano. Fonte: Folha de São Paulo. Disponível em: <<https://www1.folha.uol.com.br/cotidiano/2022/06/sao-paulo-construiu-12-milhao-de-apartamentos-em-seis-decadas.shtml>>. Acesso em 26 de junho de 2022.

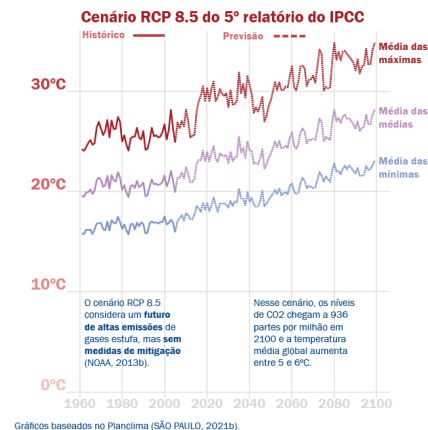
Infográficos 14 e 15: Projeta

Aumento de temperatura previsto para a cidade de São Paulo Segundo modelo HadGEM2-ES



52

No cenário sem medidas de mitigação, a média das máximas sempre ficaria acima de 30°C a partir de 2030



Infográfico 14: "Temperaturas - Projeta" 53

Descrição: Gráficos de linhas e de barras mostrando diferentes variáveis de 1961 a 2005 (registro histórico) e de 2006 a 2099 (previsão). As variáveis são temperaturas médias anuais (máximas, médias e mínimas), umidade e dias consecutivos sem chuva.

Objetivo: Na narrativa, explicitar que o aumento da temperatura média do planeta tem impactos globais. As variáveis climáticas antes mostradas por ano tendem a se intensificar.

Fonte: Dados gerados pelo CPTEC/INPE e disponibilizados na Plataforma PROJETA.

Metodologia: Os dados foram baixados a partir da API (Interface de Programação de Aplicação) da plataforma PROJETA. Foi escolhido o Cenário Climático “sudeste-BR, HADGEM2-ES” e inserido o valor de latitude e longitude para o Centro de São Paulo (latitude -23,55 e longitude -46,636). Foram escolhidas as variáveis TP2M (temperatura do ar a 2m); MXTM (temperatura máximo do ar a 2m); MNTM (temperatura mínima do ar a 2m); UR2M (umidade relativa do ar a 2m); CDD (número máximo de dias consecutivos sem chuva no ano). Em seguida, foi feito o tratamento em R, gerando os gráficos com o auxílio do pacote ggplot2.

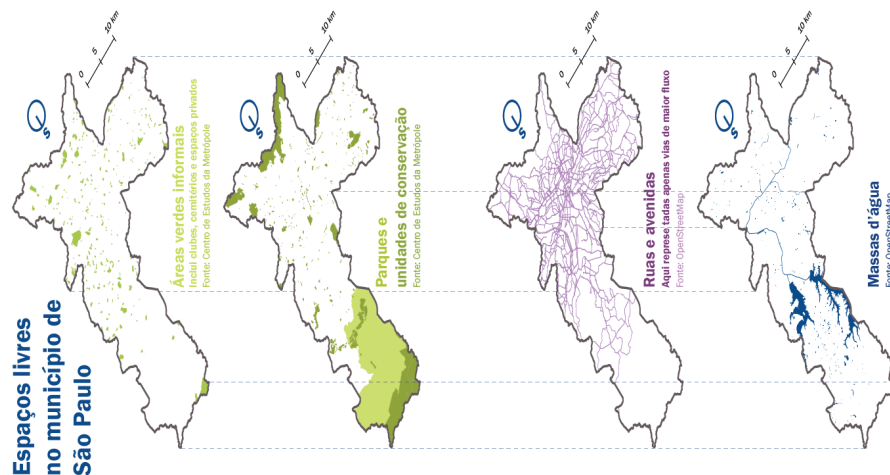
Discussão: Uma expressiva parte dos documentários e reportagens que trabalham com mudança do clima utilizam dados de previsões para o futuro, principalmente de aumento de temperatura. Este trabalho não é diferente, mas o processo de obter esses dados não é tão simples.

A partir dos gráficos do Plano de Ação Climática de São Paulo, foi feita uma leitura atenta ao documento e à bibliografia utilizada no documento. A referência, porém, era um documento do C40 (organização de quarenta cidades focada em resiliência) que não estava disponível ao público. O texto referenciava o modelo climático regional ETA_HadGEM2-ES (SÃO PAULO, 2021).

Após pesquisas sobre o modelo climático, chegou-se na plataforma PROJETA, desenvolvida pelo CPTEC/Inpe. Com o auxílio do mestrando Daniel Yoshida (FAUUSP), entendeu-se que o cenário climático mais adequado era o “05 km, sudeste-BR, HADGEM2-ES”. Os quilômetros citados são a unidade de precisão do modelo em relação às coordenadas geográficas; o sudeste-BR é relativo ao mapa utilizado para cálculo, que engloba a região Sudeste do Brasil; por fim, HADGEM2-ES é o modelo utilizado nas simulações.

Não foi possível fazer uma requisição simplificada dos dados; com a ajuda do cientista de dados Gabriel Zanlorenssi, foi feito o download utilizando uma API disponibilizada pela própria plataforma. Dez variáveis diferentes foram testadas, com resultados diversos: algumas não mostravam variações significativas, outras não tinham resultados claros o suficiente num gráfico para serem incluídas. Assim, ao final, foram selecionadas as quatro variáveis finais, escolhidas também com base no Plano de Ação Climática de São Paulo.

Infográfico 16: Espaços livres



60

Infográfico 16: "Espaços livres" 61

Descrição: Quatro mapas do município de São Paulo em isométrica mostrando espaços livres. O conjunto é conectado por linhas tracejadas, como num desenho técnico.

Objetivo: Localizar as áreas com vegetação, água e as vias da cidade. Na narrativa, fazer relação com a ideia de infraestrutura verde e azul.

Fonte: Centro de Estudos da Metrópole (áreas verdes informais, parques e unidades de conservação) e OpenStreetMap, via Geofabrik.de (vias e massas d'água).

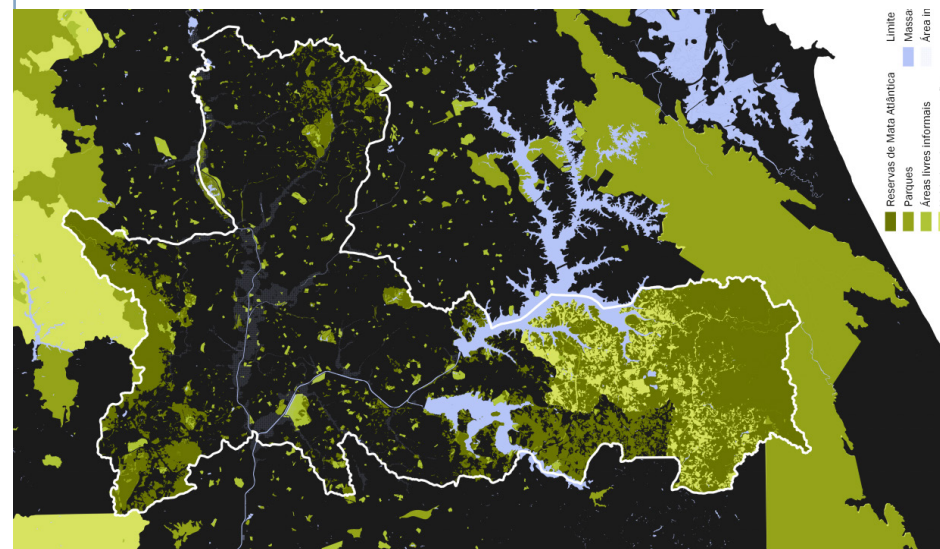
Metodologia: Os mapas foram feitos em QGIS e transformados em isométrica utilizando o Illustrator. Para melhor desempenho no programa de edição de imagem, as bases de dados foram simplificadas a partir da plataforma Mapshaper.

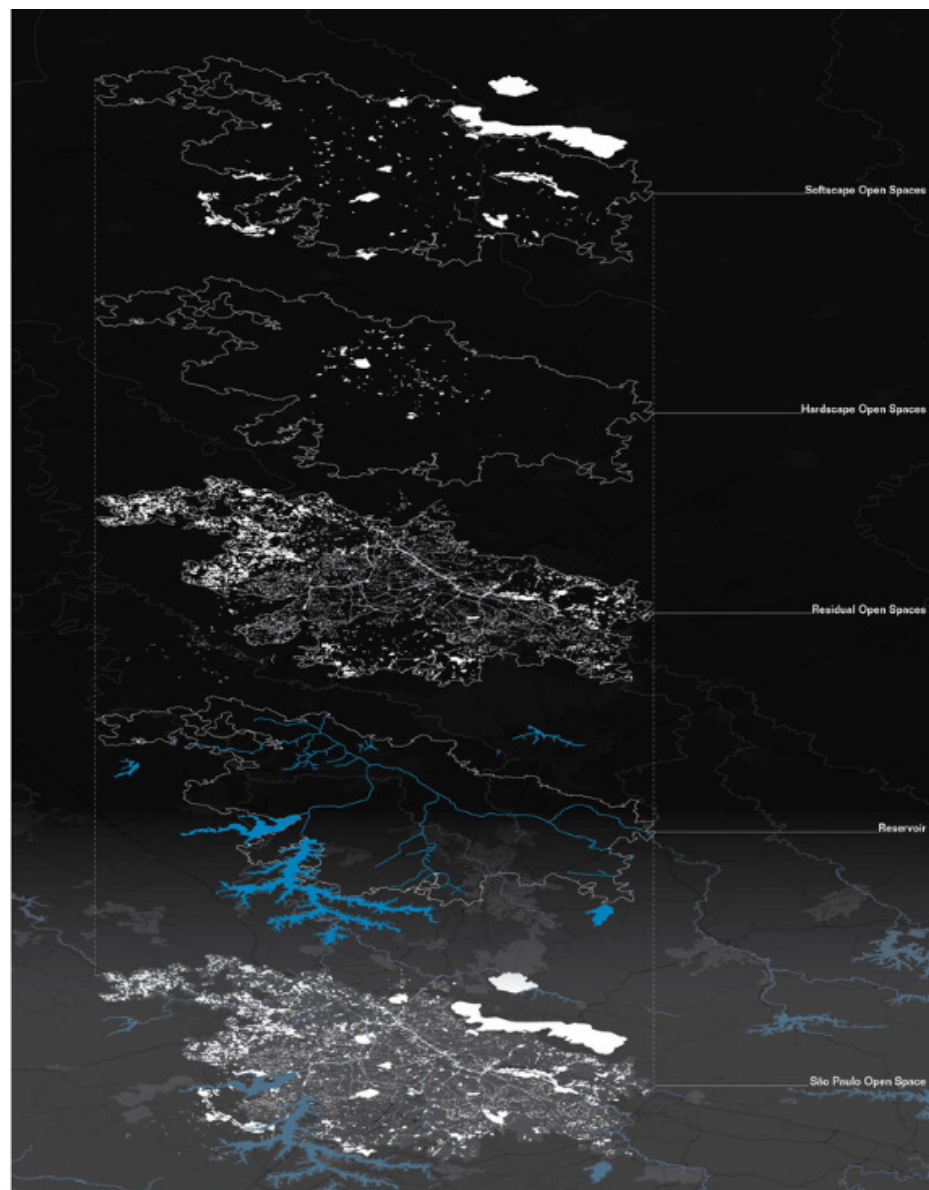
Discussão: O contato com os dados de áreas verdes do Centro de Estudos da Metrópole partiu da elaboração de materiais sobre unidades de conservação para o jornal Nexa (SOUZA; PRETTO, 2022). Eles se mostraram mais detalhados do que as informações disponíveis no Geosampa.

Também foi utilizada a base de dados do OpenStreetMap para as massas d'água. Apesar de não ser um repositório oficial do governo, se mostrou o mais abrangente e com menos falhas do que o Geosampa ou a Agência Nacional de Águas. Para complementar o sistema de espaços livres, o mesmo OpenStreetMap foi utilizado para representar as vias.

Os primeiros testes gráficos foram com fundo preto, para dar destaque às áreas verdes (figura 56). Eles foram feitos no mesmo período que o infográfico 5 (reservatórios), e contribuíram para a compreensão de que o fundo branco era um elemento unificador da linguagem visual do trabalho e deveria ser preservado. A confecção do mapa em preto também explicitou o problema da sobreposição de camadas. Qual a ordem mais correta de representação de distintas categorias de áreas verdes, se as feições se sobrepõem?

Figura 56: Versão inicial do mapa de espaços livres.
Fonte: o autor.

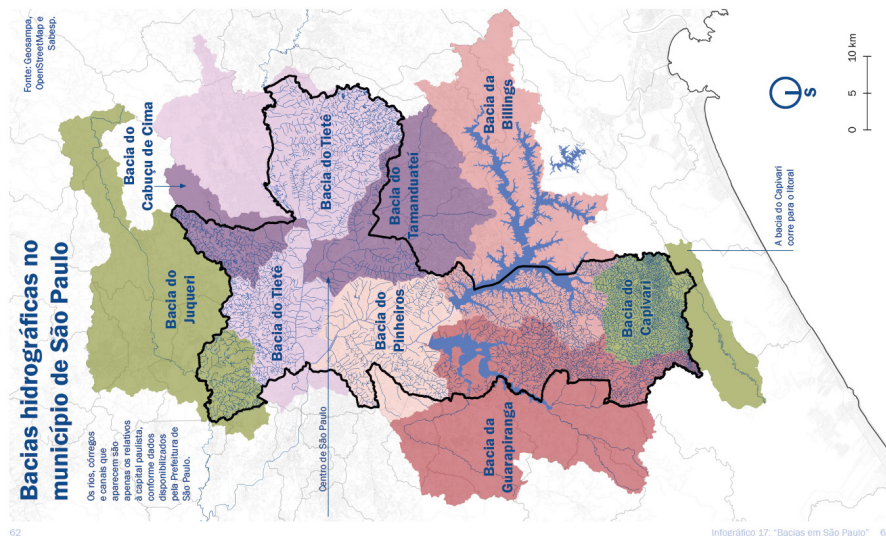




Enquanto tinha dúvidas de como proceder com esse mapa, tive contato com a biografia gráfica de São Paulo de Correa (2018). No (belíssimo) livro, o arquiteto expõe um grande número de mapas para explicar a cidade. Alguns deles colocam o mapa da região em perspectiva e trabalham várias variáveis separadamente, como no caso do próprio mapa de infraestrutura verde e azul (figura 57). As isométricas sobrepostas dão uma sensação melhor de compor um todo do que mapas planos colocados lado a lado; por isso foi o formato definido para o infográfico deste trabalho de conclusão.

Figura 57: Página da biografia gráfica de São Paulo com os espaços livres. Fonte: CORREA, 2018, via Somatic Collaborative. Disponível em: <<https://www.somatic-collaborative.com/research/sao-paulo-a-graphic-biography>>. Acesso em 26 de junho de 2022.

Infográfico 17: Bacias de São Paulo



Descrição: Mapa com os rios, represas e bacias hidrográficas de São Paulo. Os corpos d'água aparecem em azul, enquanto as bacias são representadas com tons diversos. Ao fundo, um mapa de vias.

Objetivo: Apresentar os rios, córregos e bacias hidrográficas de São Paulo. Na narrativa, contribuir para a explicação da infraestrutura azul e reforçar o grande número de corpos hídricos presentes no município.

Fonte: Geosampa (drenagem municipal e bacias hidrográficas) e OpenStreetMap (vias).

Metodologia: Trabalho feito em QGIS.

Discussão: Um trabalho sobre as águas de São Paulo não pode deixar de mostrar o grande número de rios na cidade. A Prefeitura, a partir da plataforma Geosampa, disponibiliza as informações da drenagem municipal, o mais próximo de uma "base de dados dos rios". Infelizmente, não foram encontrados dados sobre o estado de cada um (se está sobreposto por alguma via, grau de poluição, vazão etc).

A ideia inicial era mostrar apenas os rios e córregos da cidade, como no mapa proposto pelo G1 em 2017 (figura 58); porém, a grande área total do município faz com que esses corpos hídricos não expressem, sozinhos, uma narrativa. O tamanho da folha a ser impressa (menor que uma A3) também não permite detalhar muitos nomes, o que acabaria levando a escolhas arbitrárias sobre quais rios especificar.

Analisando as bases de dados do próprio Geosampa, decidiu-se pela inclusão das bacias hidrográficas. Pela sua pequena quantidade (oito) em contraste com a infinidade de rios e córregos, o mapa ficou de mais simples compreensão.

Veja o mapa dos rios e córregos de São Paulo

A capital paulista tem uma série de córregos canalizados e rios ocultos; veja qual fica mais perto da sua casa

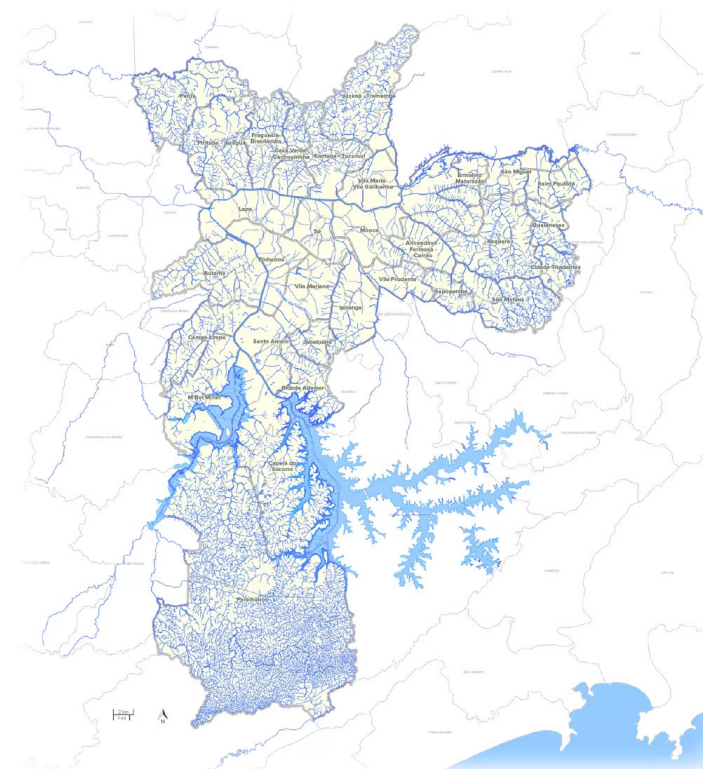
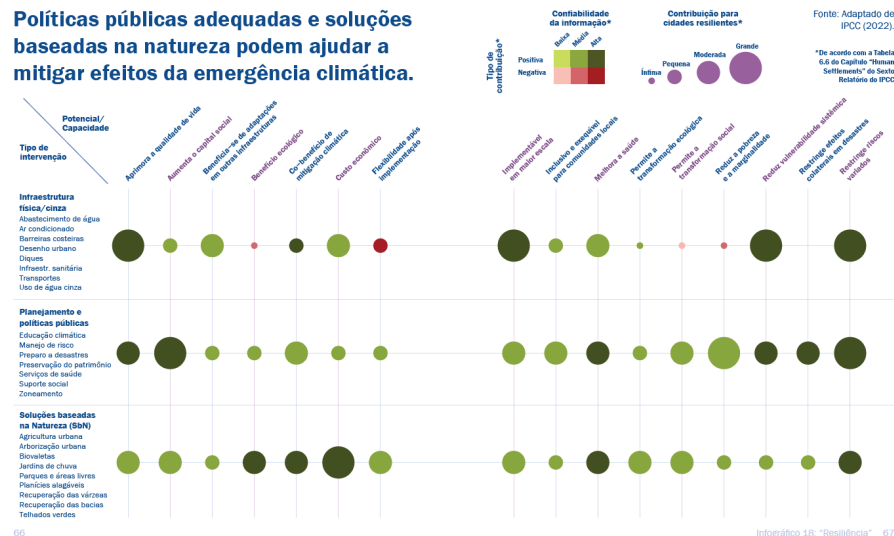


Figura 58: Mapa de rios e córregos proposto pelo G1. Fonte: G1, 2017. Disponível em: <<https://g1.globo.com/sao-paulo/rios-de-sao-paulo/noticia/mapa-mostra-rios-e-corregos-de-sao-paulo.ghtml>>. Acesso em 2 de julho de 2022.

Infográfico 18: Resiliência

Políticas públicas adequadas e soluções baseadas na natureza podem ajudar a mitigar efeitos da emergência climática.



Descrição: Tabela comparando os tipos de intervenção urbana com o potencial/capacidade de intermediar determinadas reações.

Objetivo: Reforçar a importância das adaptações urbanas frente ao cenário climático presente e futuro.

Fonte: Factsheet "Human Settlements" do AR6, IPCC.

Metodologia: Tabulação em Excel, gráfico básico em R e acabamento em Illustrator.

Discussão: A frente de trabalho essencial mais óbvia em matéria de emergência climática é a redução das emissões de gases estufa. A maioria dos tratados internacionais lida com isso, e ela é, de fato, a principal forma de mitigar esse ciclo destrutivo. Quanto antes a neutralidade em carbono global possa ser atingida, menores serão os efeitos da industrialização na atmosfera no longo prazo.

Uma segunda frente essencial é a adaptação das espécies e ambientes às novas realidades: os infográficos 9, 10 e 11 deste trabalho (chuvas no verão, tempestade e garoa e Cantareira) revelam que mudanças no clima paulistano já estão em cur-

so. As SbN (Soluções baseadas na Natureza) são aplicáveis ao ambiente rural e urbano (SEDDON, 2022), além de serem necessários investimentos em estruturas tradicionais de engenharia e em políticas públicas visando a mitigação dos problemas.

A primeira frente é indissociável de dados. Geralmente, nosso primeiro contato com a informação das emissões de gases estufa é por gráficos que quantificam esses valores, como no infográfico 13 (emissões). A segunda frente é menos óbvia de ser tabulada: qual seria a variável adequada para medir a adaptação às mudanças climáticas? É complexo encontrar métricas que unifiquem várias iniciativas diferentes.

O IPCC (Painel Intergovernamental sobre Mudanças Climáticas das Nações Unidas), responsável por concentrar várias pesquisas sobre o tema, lançou em 28 de fevereiro a segunda parte do 6º Relatório (AR6), chamada Impactos, Adaptação e Vulnerabilidade, com o capítulo 6 reservado à discussão urbana e das infraestruturas. No fact sheet (“ficha técnica”, um resumo de conceitos abordados) desse capítulo, foi disponibilizado um gráfico relacionando as formas de adaptação (infraestrutura física, planejamento e políticas públicas e SbN) com possíveis impactos e reações (figura 59).

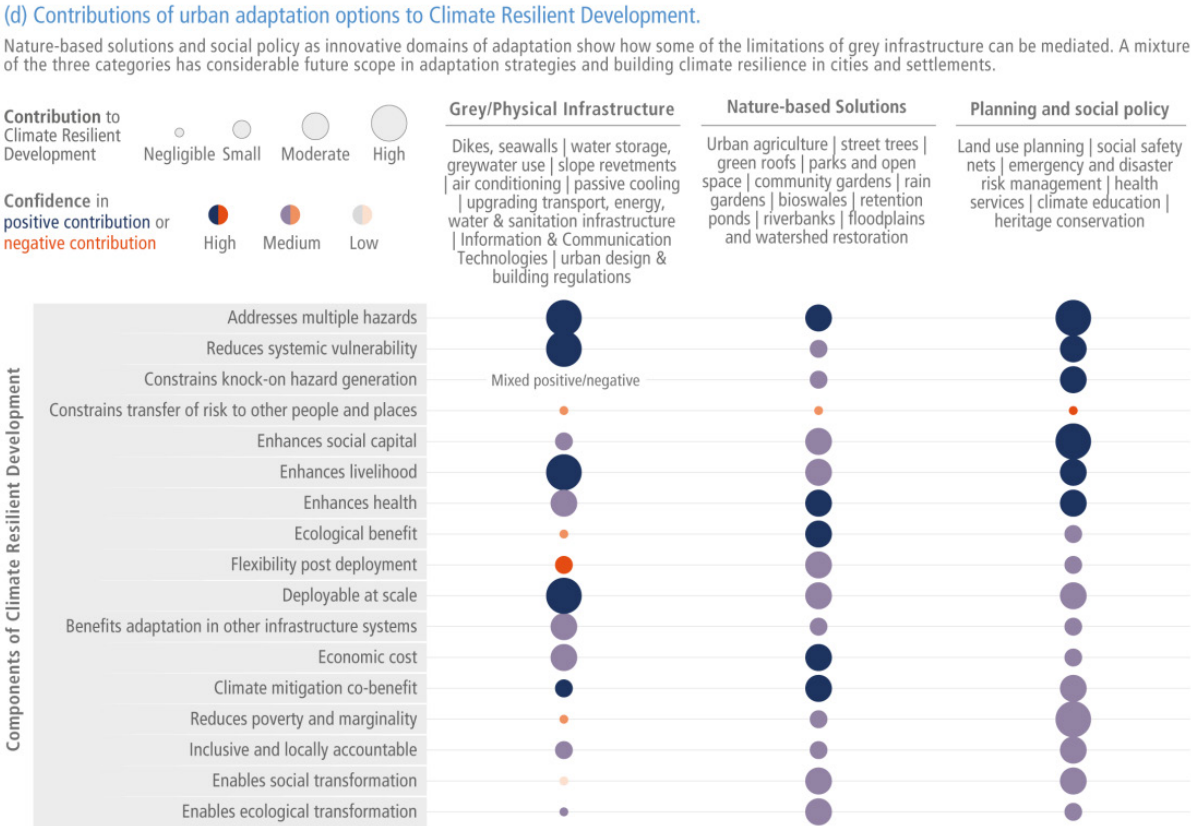
O gráfico se mostrou excelente para a narrativa deste trabalho de conclusão. Enquanto produção oficial do IPCC, é de altíssima credibilidade, apesar de apresentar as soluções em categorias amplas (um urbanista dificilmente colocaria “ar condicionado”, “melhorias na infraestrutura de transportes” e “desenho urbano” no mesmo pacote). Sua função é de dar início ao debate sobre a importância dessas iniciativas, introduzindo (por exemplo) projetos que as englobem.

Muita reflexão foi investida sobre a aparência do gráfico, mas a tabela realmente parece a melhor maneira de ilustrar a complexidade das variáveis. Foi decidido manter o formato do gráfico, alterando poucos elementos. As cores foram alteradas conforme a identidade visual do trabalho; os eixos x e y foram invertidos para se ade-

quar às páginas; optou-se por retirar a reação “constrains knock-on hazard generation” pelo interesse em manter uma representação integralmente visual, sem a opção “misto positivo/negativo”.

Muitas vezes, se um profissional anterior optou por certa visualização, é porque ela realmente é a que mais faz sentido.

Figura 59: Gráfico original sobre adaptação urbana. Fonte: IPCC, 2022. Disponível em: <https://www.ipcc.ch/report/ar6/wg2/downloads/outreach/IPCC_AR6_WGII_FactSheet_HumanSettlements.pdf>. Acesso em 22 de junho de 2022.



5. Considerações finais

5.1. Lições aprendidas

Este trabalho final de graduação não percorreu uma trajetória linear. Talvez justamente por isso tenha sido um processo de grande aprendizado pelo qual sou grato. Ainda assim, reconhecer falhas e pensar criticamente pode contribuir para que futuros trabalhos, que almejem resultados parecidos, sejam elaborados com menos “tentativa e erro”.

Deixo aqui algumas sugestões procedimentais. O termo “bibliografia complementar” vai ser utilizado como sinônimo de “bibliografia para o caderno”, em contraste com a literatura sobre visualização de dados e infografia usada no relatório.

Tenha pessoas com quem conversar sobre visualização de dados e sobre o tema escolhido.

Foi fundamental ter tido contato com o Prof. Leandro Velloso, auxiliando nas visualizações, e com a Prof.^a Denise Duarte, dando respaldo às leituras sobre mudanças climáticas. Não teria sido possível fazer o trabalho sem o auxílio de qualquer uma das partes. A dinâmica “orientador e coorientadora” foi um grande privilégio permitido pela FAUUSP; se isso não for possível, escolha uma pessoa orientadora da área que você tenha menor conhecimento e busque pessoas externas para conversar sobre os outros pontos, mantendo a pessoa orientadora informada.

Delimite muito bem o seu tema.

Um trabalho de conclusão não precisa mudar o mundo: no meu caso, não vai resolver a crise climática. Justamente por isso, quanto mais bem delimitado o tema, maiores chances de abordá-lo com a devida complexidade. Se a leitura da bibliografia faz parecer que o tema é grandioso demais para seu momento de vida, converse com a pessoa orientadora sobre quais subtópicos que mais te interessam, para assim restringir.

Seja fiel aos prazos máximos de leitura da bibliografia complementar.

O tema escolhido para este trabalho era extremamente abrangente e difícil de abarcar em uma parcela de semestre. Enquanto lia relatórios e artigos sobre emergência climática, fiquei paralisado pensando que o tema era muito complexo; entendi só após semanas que não é preciso ser um PhD em determinado tema para compor uma narrativa baseada em dados. Respeitar prazos máximos de leitura ajuda a reconhecer os limites na aquisição de conhecimento.

Leia a bibliografia complementar pensando na estrutura do roteiro.

O principal objetivo da literatura é respaldar o conjunto dos infográficos. Fazer essa leitura já pensando em como vai utilizá-la no futuro facilita muito a escrita do roteiro. Ao ler um texto, não apenas grife partes importantes, mas catalogue em que momento da narrativa ele pode ser utilizado. A estrutura, em si, pode nascer dessa leitura estratégica.

Comece a explorar suas bases de dados o quanto antes. Visualização de dados só existe caso existam dados. Um erro que cometi foi pensar que, por catalogar e fazer download de inúmeros arquivos, já tinha dados suficientes. Um fato, porém, é que quantidade difere de qualidade: um dataset com descrição muito interessante pode ser pouco representativo, ter muitas falhas ou não corresponder à informação precisa que parecia conter. Entender suas bases de dados ainda cedo evita desilusões no futuro: tente gerar gráficos ou tabelas simples, só para ter uma noção das informações.

Numa guerra de três frentes (narrativa, trabalho com dados e design), as batalhas não precisam acontecer numa ordem pré-definida. A ordem natural deste trabalho me parecia ser “roteiro, dados e infografia”; quando tive dificuldades com o roteiro, passei semanas paralisado sem avançar. Com a insistência do Prof. Leandro Velloso, passei a explorar os dados disponíveis mesmo sem o roteiro pronto. Foi uma sábia decisão: os infográficos produzidos naturalmente geravam lacunas que precisavam ser preenchidas com o roteiro. Se alguma das três frentes do trabalho estiver se mostrando muito difícil, procure avançar nas outras.

O todo é maior que a soma das partes. Se alguma base de dados é muito trabalhosa, pare e reflita quão essencial ela é para o conjunto da obra. Alguns dos datasets mais trabalhosos usados neste trabalho de conclusão, que envolveram mais de duas semanas de angústias, resultaram em apenas um infográfico, no máximo, dois. Se algum deles estiver travando seu fluxo de trabalho, deixe ele de lado por um tempo e volte a trabalhar nele após arejar a mente trabalhando em outros.

Peça ajuda a seu orientador ou orientadora em momentos de paralisia. Talvez o maior erro que cometi durante esse processo tenha sido demorar para comentar com a Prof.^a Denise Duarte que estava tendo dificuldade com a literatura sobre mudanças climáticas, o que estava travando o restante do trabalho. Eu tinha vergonha de chegar numa conversa de mãos vazias, pensando que “na próxima semana vou ter vários infográficos para mostrar” (o que não acontecia). Quando finalmente quebrei essa barreira mental, tive uma conversa importantíssima com a professora, que ajudou a botar o trabalho “nos eixos”.

Procure não deixar de fazer atividade física durante o trabalho de conclusão.

5.2. Conclusão

O objetivo deste trabalho era elaborar uma narrativa de jornalismo de dados, baseado em infografia, com o respaldo de conhecimentos pertinentes à arquitetura e ao urbanismo. Sendo literal à proposta, entendo que ela foi parcialmente concretizada: compreendo que a abordagem de jornalismo de dados não se mostrou a mais adequada, já que os dados serviram como respaldo a uma narrativa definida através de pesquisa acadêmica, não necessariamente por meio de uma exploração heroica de bases de dados.

Ainda assim, entendo que foi realizado um grande exercício em muitas frentes, do qual tenho orgulho. O caderno que acompanha esta monografia é fruto de intenso trabalho de pesquisa, reflexão, tratamento de dados e produção visual.

A diversidade de tarefas envolvidas pode ter levado algumas a serem levadas com mais atenção do que outras. Entendo que, num mundo ideal, a parte final da narrativa poderia elaborar mais sobre as políticas públicas e soluções de adaptação possíveis; gostaria também de ter explorado dados do orçamento público municipal.

Ainda assim, uma das lições aprendidas se sobrepõe a essas inseguranças: **o todo é maior que a soma das partes.** Não se pode investir em elementos específicos em detrimento do conjunto.

Um efeito colateral deste trabalho foi o despertar de um profundo interesse por entender melhor variáveis e dinâmicas relacionadas à emergência climática, não apenas os gráficos e mapas resultantes delas. O tema é complexo e gerou muito nervosismo no início da elaboração do caderno, mas se mostrou muito interessante e com vários pormenores aplicáveis à discussão em arquitetura e urbanismo.

6. Referências bibliográficas

99% INVISIBLE. Florence Nightingale: Data viz pioneer. [Locução de]: Roman Mars, Tim Harford. **99% Invisible**, 2 de março de 2021. Disponível em <<https://99percentinvisible.org/episode/florence-nightingale-data-viz-pioneer/>>. Acesso em 21 de novembro de 2021.

ARAÚJO, Gabriela Momberg. **Dados à mesa: visualização de sistemas alimentares no território**. Trabalho final de graduação (Arquitetura e Urbanismo) — Universidade de São Paulo (USP). São Paulo, 2021.

BENEVOLO, Leonardo. **História da Cidade**. São Paulo: Perspectiva, 2011.

BERINATO, Scott. **Good charts workbook: tips, tools, and exercises for making better data visualizations**. Boston, Massachusetts: Harvard Business Review Press, 2019.

BOUNERGU, Liliana. Data Journalism in Perspective. In: **The Data Journalism Handbook**. P. 17-22. Organizadores: GRAY, Jonathan et al. Sebastopol: O'Reilly, 2012. Disponível em: <https://datajournalismhandbook.org/uploads/first_book/DataJournalismHandbook-2012.pdf>. Acesso em 29 de novembro de 2021.

BRADSHAW, Paul. What is Data Journalism? In: **The Data Journalism Handbook**. P. 2-3. Organizadores: GRAY, Jonathan et al. Sebastopol: O'Reilly, 2012. Disponível em: <<https://datajournalismhandbook.org/>

[uploads/first_book/DataJournalismHandbook-2012.pdf](https://datajournalismhandbook.org/uploads/first_book/DataJournalismHandbook-2012.pdf)>. Acesso em 29 de novembro de 2021.

BYRON, Lee; WATTENBERG, Martin. Stacked Graphs – Geometry & Aesthetics. **IEEE Transactions on Visualization and Computer Graphics**, v. 14, n. 6, p. 1245–1252, nov. 2008.

CHEN, Chun-Houn. **Handbook of Data Visualization**. Nova York: Springer, 2008.

CLARKE, Keir. **The New Racial Dot Maps of America**. 14 de abril de 2022. Disponível em: <<https://google-mapsmania.blogspot.com/2022/04/the-new-racial-dot-maps-of-america.html>>. Acesso em 29 de junho de 2022.

CODDINGTON, Mark. Clarifying Journalism's Quantitative Turn. **Digital Journalism**, 7 de novembro de 2014. Disponível em: <<https://www.tandfonline.com/doi/full/10.1080/21670811.2014.976400>>. Acesso em 29 de novembro de 2021.

DI BARTOLOMEO, Marco; HU, Yifan. There is More to Streamgraphs than Movies: Better Aesthetics via Ordering and Lassoing. **Computer Graphics Forum**, v. 35, n. 3, p. 341–350, jun. 2016.

DICIONÁRIO MICHAELIS. **Infografia**. n.d. Disponível em: <<https://michaelis.uol.com.br/busca?r=0&f=0&t=0&pa>

[lavra=infografia](https://michaelis.uol.com.br/busca?r=0&f=0&t=0&pa)>. Acesso em 29 de novembro de 2021.

EMPLASA. **Expansão da área urbana da Região Metropolitana de São Paulo (versão preliminar)**. Relatório. São Paulo, 2007. Disponível em: <<http://www.igc.sp.gov.br/produtos/arquivos/ExpansaoDaAreaUrbanizadaDaRmSPAt2002.pdf>>. Acesso em 26 de junho de 2022.

EGAN, Dan. The climate crisis haunts Chicago's future. A Battle Between a Great City and a Great Lake. **The New York Times**, 7 de julho de 2021. Disponível em: <<https://www.nytimes.com/interactive/2021/07/07/climate/chicago-river-lake-michigan.html>>. Acesso em 19 de novembro de 2021.

FRIENDLY, Michael. A Brief History of the Mosaic Display. **Journal of Computational and Graphical Statistics**, v. 11, n. 1, p. 89–107, mar. 2002.

FRIENDLY, Michael; WAINER, Howard. **A History of Data Visualization and Graphic Communication**. Cambridge: Harvard University Press, 2021.

HARFORD, Tim. **The Data Detective: Ten Easy Rules to Make Sense of Statistics**. Nova York: Penguin Publishing Group, 2021.

HIGHWOOD, Ellie. **#climatechange: The global warming blanket**. 12 de junho de 2017. Disponível em: <<https://elliehighwood.com/2017/06/12/climatechange-crochet-the-global-warming-blanket/>>. Acesso em 30 de junho de 2022.

com/2017/06/12/climatechange-crochet-the-global-warming-blanket/>. Acesso em 30 de junho de 2022.

HOLOVATY, Adrian. **A fundamental way newspaper sites need to change**. 6 de setembro de 2006. Disponível em: <<http://www.holovaty.com/writing/fundamental-change/>>. Acesso em 29 de novembro de 2021.

IABLONOVSKI, Guilherme M. How to Download Uber's Hexagonal Grid with Python. **Towards Data Science**, 6 de junho de 2022. Disponível em: <<https://towardsdatascience.com/how-to-download-ubers-hexagonal-grid-with-python-3140fe95e19a>>. Acesso em 25 de junho de 2022.

KALACH, Alberto. **México: Ciudad Futura**. Cidade do México: Editorial RM, 2010.

KNAFLIC, Cole Nussbaumer. **Storytelling com dados**. Rio de Janeiro: Alta Books, 2019.

LACERDA JÚNIOR, Aécio Flávio de Souza. **Habitação de Interesse Social: Jardim Edith da favela ao conjunto residencial**. Dissertação (mestrado em Arquitetura e Urbanismo) - Universidade de São Judas Tadeu. São Paulo, 2016.

LEHNER, Bernhard; GRILL, Günther. Global river hydrography and network routing: baseline data and new approaches to study the world's large river systems: GLOBAL RIVER HYDRO-

GRAPHY AND NETWORK ROUTING. **Hydrological Processes**, v. 27, n. 15, p. 2171–2186, 15 jul. 2013.

LITTLE, Stephen. **Ismos: para entender a arte**. São Paulo: Globo, 2010.

MARENGO, J. A. et al. Trends in extreme rainfall and hydrogeometeorological disasters in the Metropolitan Area of São Paulo: a review. **Annals of the New York Academy of Sciences**, v. 1472, n. 1, p. 5–20, jul. 2020.

MARINO, Aluizio, et al. Vacinação avançada de forma desigual em São Paulo. **Labcidade**, 21 de julho de 2021. Disponível em: <<http://www.labcidade.fau.usp.br/vacinacao-avanca-de-forma-desigual-em-sao-paulo/>>. Acesso em 26 de junho de 2022.

MORAES, Ary. **Infografia: História e projeto**. São Paulo: Blucher, 2013.

MUNSELL, Mike. Who really invented the climate stripes? **Canary Media**, 3 de junho de 2022. Disponível em: <<https://www.canarymedia.com/articles/climate-crisis/who-really-invented-the-climate-stripes>>. Acesso em 30 de junho de 2022.

O’MURCHU, Cynthia. An Essential Part of the Journalists’ Toolkit. In: **The Data Journalism Handbook**. P. 9-10. Organizadores: GRAY, Jonathan et al. Sebastopol: O’Reilly, 2012. Disponível em <<https://datajournalismhandbook.org/uploads/>

first_book/DataJournalismHandbook-2012.pdf>. Acesso em 29 de novembro de 2021.

PALMEIRA, Eric. **DataSP: uma releitura dos dados públicos da cidade de São Paulo**. Trabalho final de graduação (Arquitetura e Urbanismo) - Universidade de São Paulo (USP). São Paulo, 2020.

PEREIRA, Rafael Henrique Moraes et al. Desigualdades socioespaciais de acesso a oportunidades nas cidades brasileiras, 2019. **Texto para Discussão Ipea**, v. 2535, janeiro de 2020. Disponível em: <<http://repositorio.ipea.gov.br/handle/11058/9586>>. Acesso em 26 de junho de 2022.

PHINEAS. Who is this Sankey guy? **Sankey Diagrams**, 12 de abril de 2007. Disponível em <<https://www.sankey-diagrams.com/who-is-this-sankey-guy/>>. Acesso em 30 de maio de 2022.

PLUNZ, Richard; SUTTO, Maria Paola. **Urban Climate Change Crossroads**. p. 65-76. Nova York: Columbia University, 2008.

RAISZ, E. The Rectangular Statistical Cartogram. **Geographical Review**, v. 24, n. 2, p. 292, abril de 1934.

ROGERS, Simon. John Snow’s data journalism: the cholera map that changed the world. **The Guardian**, 15 de março de 2013. Disponível em: <<https://www.theguardian.com/news/>

datablog/2013/mar/15/john-snow-cholera-map>. Acesso em 28 de junho de 2022.

ROYÉ, Dominic. How to create ‘Warming Stripes’ in R. **Dr. Dominic Royé**, 5 de dezembro de 2018. Disponível em: <<https://dominicroye.github.io/en/2018/how-to-create-warming-stripes-in-r/>>. Acesso em 3 de junho de 2022.

SÃO PAULO, Prefeitura do município de. **Plano de ação climática do município de São Paulo 2020–2050**. São Paulo, 2021.

SANTOS, Fábio Alexandre dos. **Domando Águas: salubridade e ocupação do espaço na cidade de São Paulo, 1875–1930**. São Paulo: Alameda, 2011.

SCHEIBEL, Willy et al. A Taxonomy of Treemap Visualization Techniques. **11th International Conference On Information Visualization Theory And Applications**. Valletta, Malta: SCITEPRESS - Science and Technology Publications, 2020. Disponível em: <<http://www.scitepress.org/DigitalLibrary/Link.aspx?-doi=10.5220/0009153902730280>>. Acesso em 28 de junho de 2022.

SEDDON, Nathalie. Harnessing the potential of nature-based solutions for mitigating and adapting to climate change. **Science**, v. 376, n. 6600, p. 1410–1416, 24 de junho de 2022.

SEDG. **What is information design?** n.d. Disponível em: <<https://segd.org/what-information-design>>. Acesso em 27 de novembro de 2021.

SOJO, Carlos Abreu. **La Infografia Periodística**. Caracas: Fondo Editorial de Humanidades y Educación, 2000.

SOUZA, Caroline; PRETTO, Nicholas. O que são e onde ficam as unidades de conservação do Brasil. **Nexo Políticas Públicas**, 27 de junho de 2022. Disponível em: <<https://pp.nexojornal.com.br/Dados/2022/06/27/O-que-s%C3%A3o-e-onde-ficam-as-unidades-de-conserva%C3%A7%C3%A3o-do-Brasil>>. Acesso em 2 de julho de 2022.

STEELE, Julie e ILIINSKY, Noah. **Beautiful Visualization: Looking at data through the eyes of experts**. Cambridge: Safari Books, 2010.

TEIXEIRA, Tattiana. A presença da infografia no jornalismo brasileiro — proposta de tipologia e classificação como gênero jornalístico a partir de um estudo de caso. **Revista Fronteiras**, IX(2), mai/ago 2007. Disponível em <<http://revistas.unisinos.br/index.php/fronteiras/article/view/5847>>. Acesso em 29 de novembro de 2021.

TEIXEIRA, Tattiana. **Infografia e Jornalismo**. Salvador: EDUFBA, 2010. Disponível em <<https://repositorio>

ufba.br/ri/handle/ri/20642>. Acesso em 29 de novembro de 2021.

TRÄSEL, Marcelo Ruschel. **Entrevistando planilhas: estudo das crenças e do ethos de um grupo de profissionais de jornalismo guiado por dados no Brasil**. Tese de doutorado em Comunicação Social. Porto Alegre: PUCRS, 2014. Disponível em <<http://tede2.pucrs.br/tede2/handle/tede/4590>>. Acesso em 29 de novembro de 2021.

TUFTE, Edward. **Envisioning information**. Cheshire: Graphics Press, 1990.

TUFTE, Edward. **The Visual Display of Quantitative Information**. Cheshire: Graphics Press, 2007.

VASCONCELLOS, Fábio et al. Cinco categorias de Jornalismo de Dados ou uma proposta para problematizar o Jornalismo a partir de dados no Brasil. **II Seminário de Pesquisa em Jornalismo Investigativo**, Universidade Anhembi-Morumbi, 2015. Disponível em: <https://projetos.abraji.org.br/seminario/PDF/2/carolina_bittencourt_fabio_vasconcellos_leonardo_mancini.pdf>. Acesso em 29 de novembro de 2021.

WARE, Colin. **Information visualization: perception for design**. San Francisco: Morgan Kaufman Publishers, 2004.

WICKHAM, Hadley e GROLEMUND, Garrett. **R for Data Science**. 2017. Disponível em <<https://r4ds.had.co.nz/index.html>>. Acesso em 28 de novembro de 2021.

