

Por que a visualização de dados falha se não pensamos sobre seu propósito

Alberto Cairo¹

Resumo: A maioria das pessoas que cria gráficos ou mapas regularmente para apresentações de slides, artigos, documentos, relatórios e afins não aprende sobre a visualização em um ambiente formal. Elas nem mesmo entendem elementos básicos da visualização, como a codificação, a propriedade visual que muda em relação aos dados que estão sendo representados. O que realmente fazemos ao projetar uma visualização de dados são números de mapas e as propriedades dos objetos, geralmente formas geométricas. O objetivo deste artigo é apresentar exemplos de como a visualização deve guiar a escolha da codificação. Quando elaborados corretamente, gráficos e mapas facilitam o conhecimento e melhoram a compreensão do significado dos dados pelos leitores.

Palavras-chave: Visualização; Codificação; Dados; Design

Abstract: Most people who create charts or maps regularly for slideshows, articles, documents, business reports and the like do not learn about visualization in a formal setting. They often do not even understand the basic elements of visualization, such as encoding, which is the visual property that changes in relation to the data that is being represented. What we really do when designing a data visualization is map numbers onto the properties of objects, usually geometric shapes. The purpose of this article is to provide examples of how the visualization should guide the choice of encoding. When done correctly, graphs and maps facilitate knowledge and improve readers' grasp of the meaning of data.

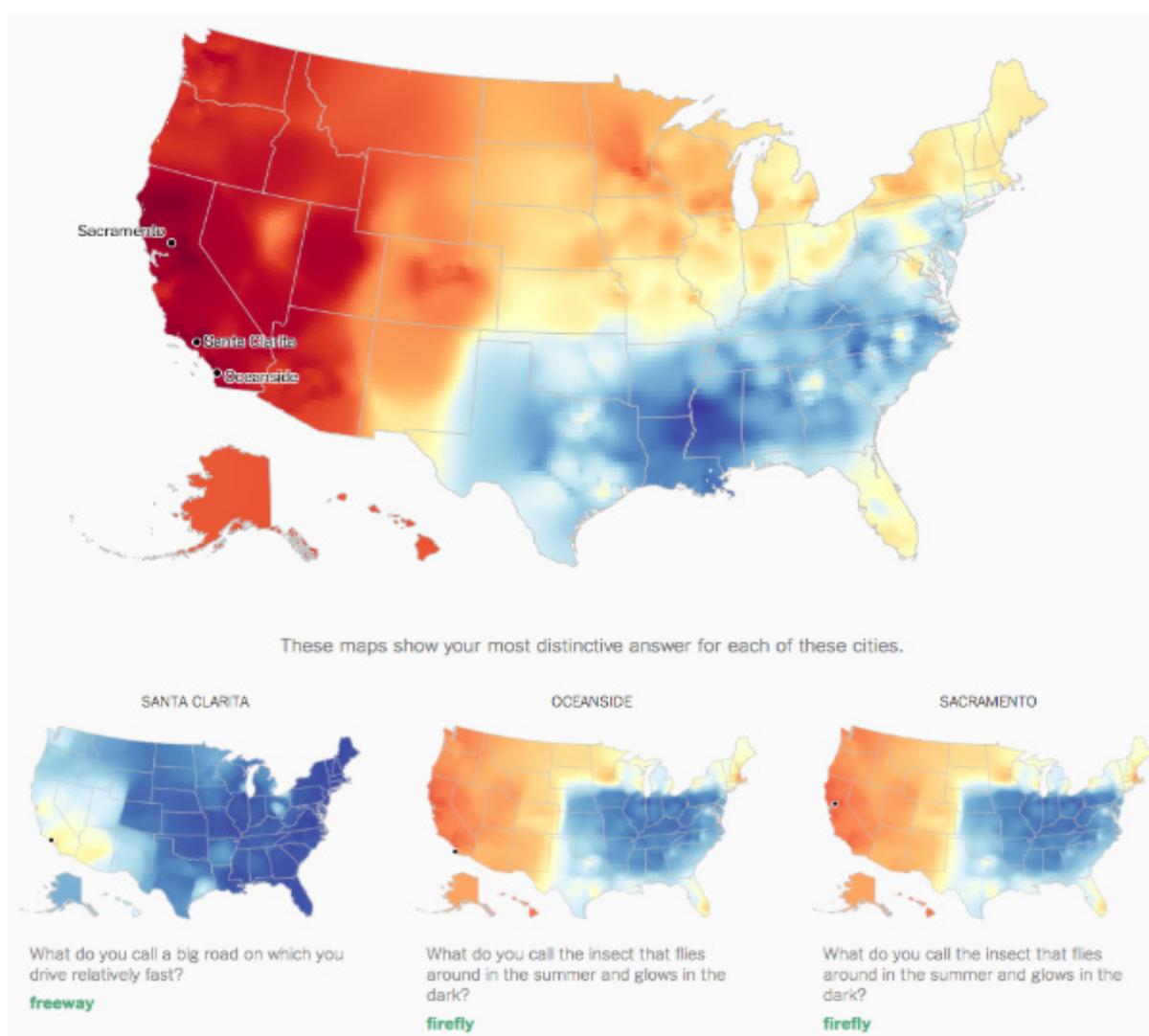
Keywords: Visualization; Encoding; Data; Design.

¹ Prof. Dr. pela Universitat Oberta de Catalunya (Barcelona, Espanha), titular da Cátedra Knight de Jornalismo Visual na Universidade de Miami, autor de *How Charts Lie* (W. W. Norton & Company, 2019). E-mail: a.cairo@miami.edu

A questão da visualização de dados

A popularidade da visualização de dados, a representação de informações por meio de gráficos, tabelas e mapas, aumentou drasticamente nas últimas duas décadas graças à crescente disponibilidade de um conjunto de informação em formatos utilizáveis em fontes de governo, não-governamental ou em plataformas corporativas, além de ferramentas de software que, em muitos casos são de livre acesso ou em código aberto. A publicação de notícias em jornais de grande ou pequena circulação contribuiu para esta explosão no uso dos gráficos de dados e sua popularidade junto aos leitores tem sido comprovada repetidamente. Um registro para corroborar este ponto: o conteúdo com o maior número de leituras publicado até hoje na edição online do *The New York Times*, o jornal mais influente nos Estados Unidos e um dos mais influentes no mundo, é uma visualização de dados intitulada “o mapa do dialeto”.

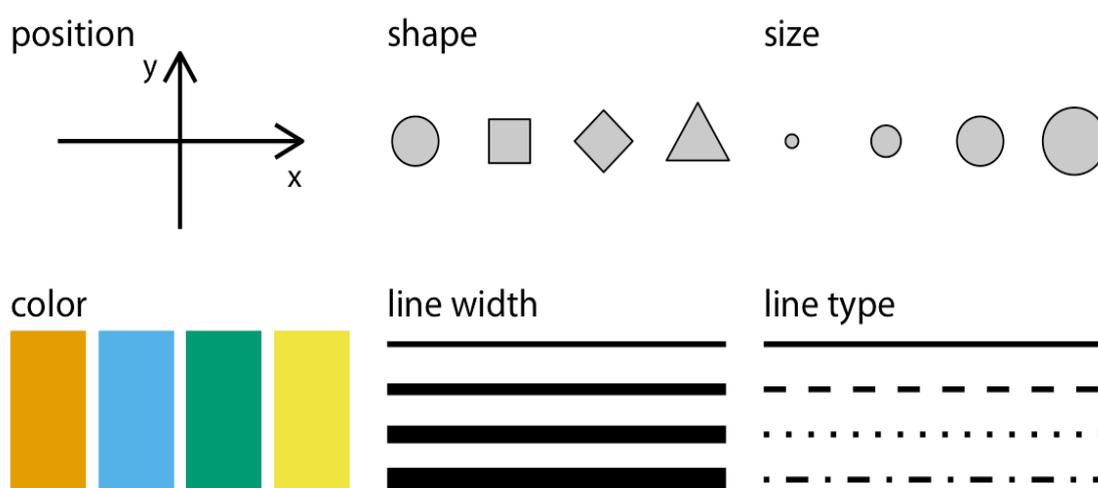
Figura 1: O Mapa Dialeto.



Fonte: The New York Times

A crescente popularidade da visualização de dados não foi condizente com atenção equivalente aos princípios elementares que devem orientar o seu design. A maioria das pessoas que criam gráficos ou mapas regularmente para apresentações de slides, artigos, documentos, relatórios de negócios e similares não aprende a visualização formalmente. Elas muitas vezes nem sequer entendem os princípios mais básicos, como a noção de codificação: o que realmente fazemos ao projetar a visualização de dados é mapear números para propriedades de objetos, geralmente formas geométricas ou “GeoMS”. O que muda em relação aos números representados é a propriedade da codificação.

Figura 2: Codificações comuns na visualização de dados



Fonte: Serial Mentor (2019)

Pense em um gráfico de barras: as “geoms” são constituídas por uma série de barras; a codificação – propriedade que varia de acordo com os números que queremos representar – é a altura ou o comprimento dessas barras. Imagine o exemplo de um mapa com múltiplas cores: a codificação é cor ou sombra: se estamos representando taxas de desemprego na Espanha, por exemplo, quanto maior o desemprego em uma província, mais escura será a tonalidade da cor de determinadas províncias.

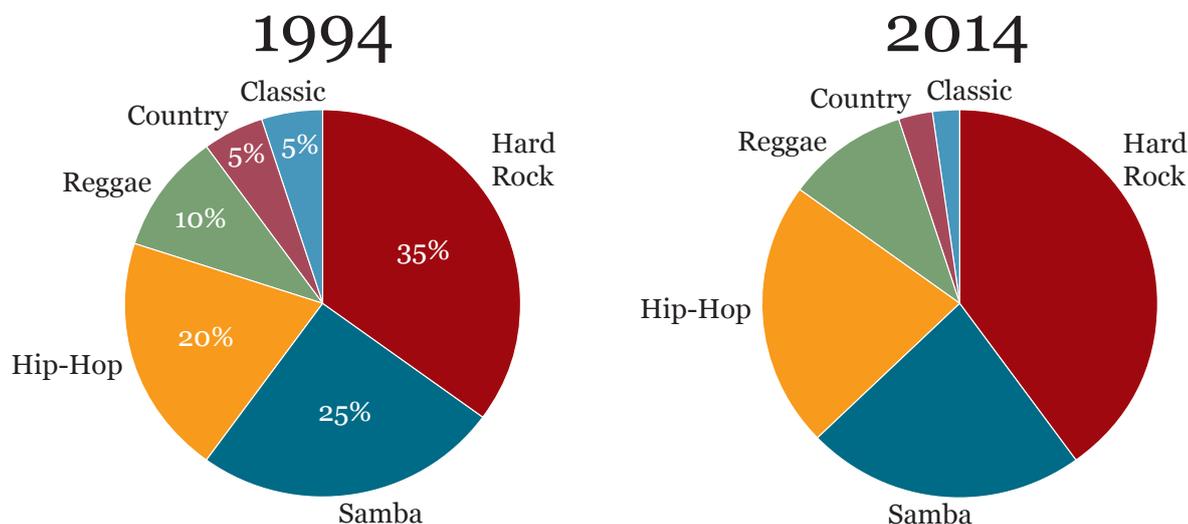
A escolha de codificações não é um processo arbitrário ou baseado apenas em preferências pessoais. A seleção deve depender do propósito da visualização. Sempre que projetamos um gráfico, precisamos nos perguntar: o que o meu leitor precisa ver neste gráfico? Que tipo de informações eu quero fornecer? Como exemplos aqui estão os dois gráficos da Figura 3, baseados em dados simulados de gêneros musicais preferidos por estudantes em dois períodos distintos. O objetivo desses gráficos está explícito no seu título: comparar a porcentagem de estudantes que escolheram seus gêneros favoritos de música em 1994 e em 2014.

Figura 3: Da arte verdadeira/ From The Truthful Art

How Music Preferences Have Changed in Two Decades

Music styles preferred by University of Miami students. Survey based on interviews with 1,000 students.

SOURCE: WishfulThinkingData Inc.



Fonte: Elaboração Própria (2019)

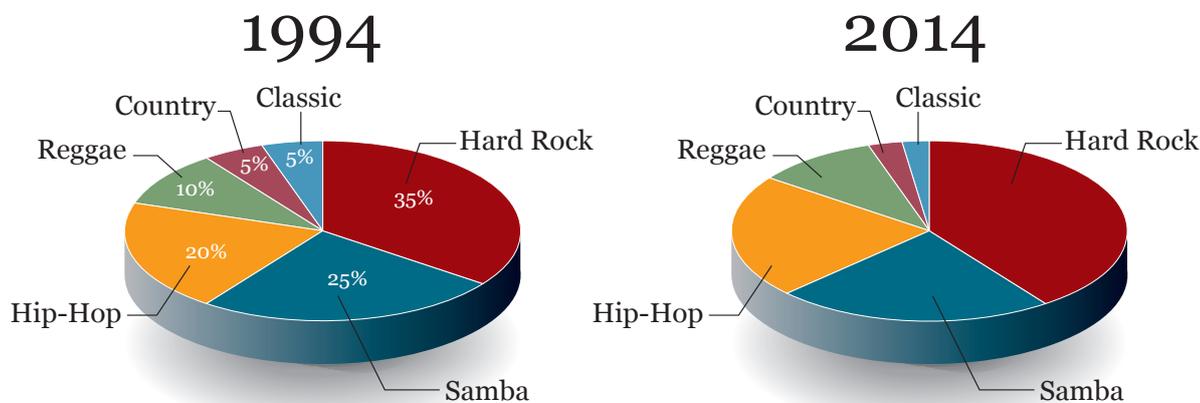
As comparações se tornam ainda mais difíceis se os gráficos forem produzidos no formato de pizza tridimensional (Figura 4), um pecado ainda muito comum nas apresentações. Aqui uma observação fundamental: nunca devemos produzir qualquer gráfico estático ou gráficos 3D ou adicionar perspectiva.

Figura 4: Da arte verdadeira/ From The Truthful Art

How Music Preferences Have Changed in Two Decades

Music styles preferred by University of Miami students. Survey based on interviews with 1,000 students.

SOURCE: WishfulThinkingData Inc.



Fonte: Elaboração própria (2019)

O problema neste exemplo é que os gráficos de pizza usam o ângulo e a área como codificações e estes não são muito bons quando a finalidade de uma visualização é representar mudança ou permitir que os leitores comparem números. A área ou o ângulo podem ter codificações apropriadas em outros casos.

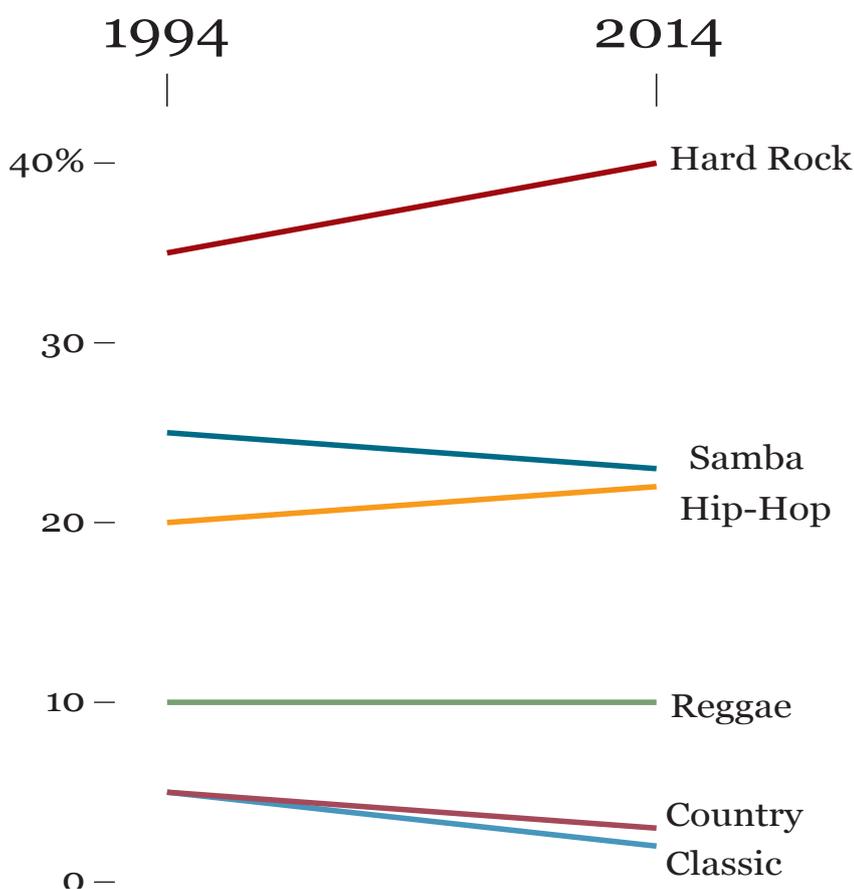
Outro exemplo é quando descrevemos os mesmos dados de maneira completamente diferente, usando a posição em escalas verticais e a inclinação da linha como codificações (Figura 5). Hoje é fácil ver os gêneros musicais que se tornaram mais ou menos populares em 1994 e em 2014. O aprendizado principal que se deduz deste exemplo é simples: se a finalidade da visualização é realmente comparar e revelar mudanças, devemos mostrar a comparação e a alteração, não as partes de um todo, como o fazem os gráficos na forma de pizza.

Figura 5: Da arte verdadeira/ From The Truthful Art.

How Music Preferences Have Changed in Two Decades

Music styles preferred by University of Miami students.
Survey based on interviews with 1,000 students.

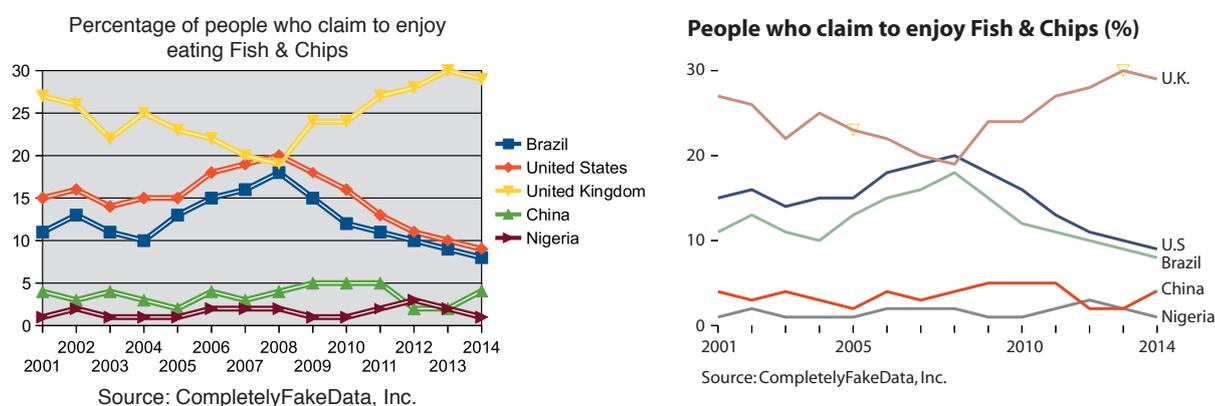
SOURCE: WishfulThinkingData Inc.



Fonte: Elaboração própria (2019)

A codificação é apenas um dos vários fatores a serem considerados ao projetar gráficos, tabelas e mapas. Cientistas, estatísticos, engenheiros, analistas de dados e outros estudiosos em campos que enfatizam os métodos quantitativos muitas vezes esquecem que o design visual e a estética são essenciais na visualização dos dados, particularmente quando é preciso apresentá-los para o público em geral. Os dois gráficos na Figura 6 fazem parte do livro *The Truthful Art* (2016). Qual deles parece mais elegante? Qual deles você se sentiria interessado em ler primeiro? Provavelmente seria o que está à direita, que demandou um tempo maior para ajustes de cores, tipo e layout.

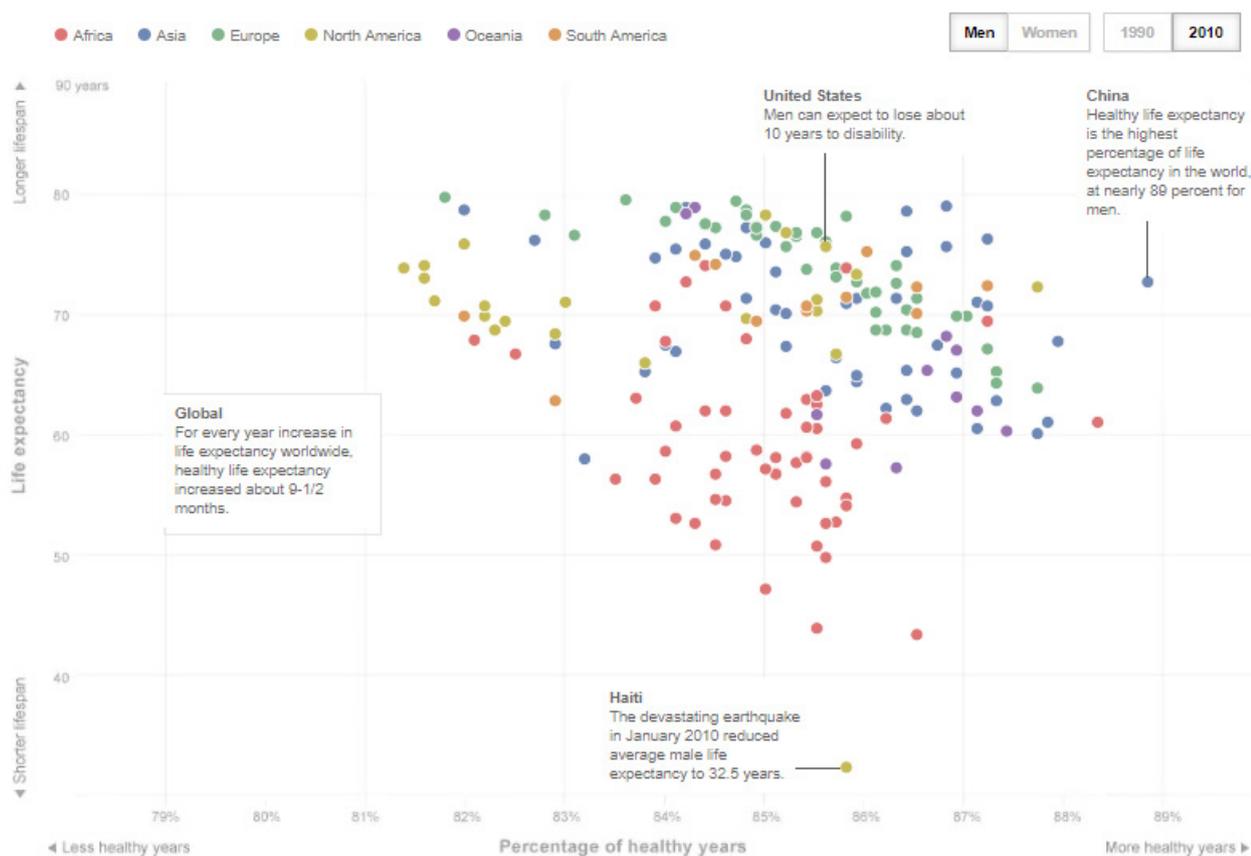
Figura 6: Da arte verdadeira/ From *The Truthful Art*.



Fonte: Elaboração própria (2019)

O gráfico à esquerda, pesado e sobrecarregado, é o resultado de deixar uma ferramenta de software, neste caso o Excel, tomar decisões de design em nosso lugar. Aprender design visual é como aprender a escrever bem: não deve ser considerado a posteriori, porque pode beneficiar a todos. Felizmente, hoje há abundância de recursos educativos sobre como ser um bom designer de dados. Finalmente, precisamos prestar atenção não apenas a codificações ou recursos de design, mas também ao texto que os acompanha: títulos, cabeçalhos, explicações, rótulos e assim por diante. Na visualização, identificamos esses elementos como “camadas de anotação”. Isso está presente no gráfico interativo do jornal *The Washington Post* a seguir (Figura 7).

Figura 7: The Washington Post



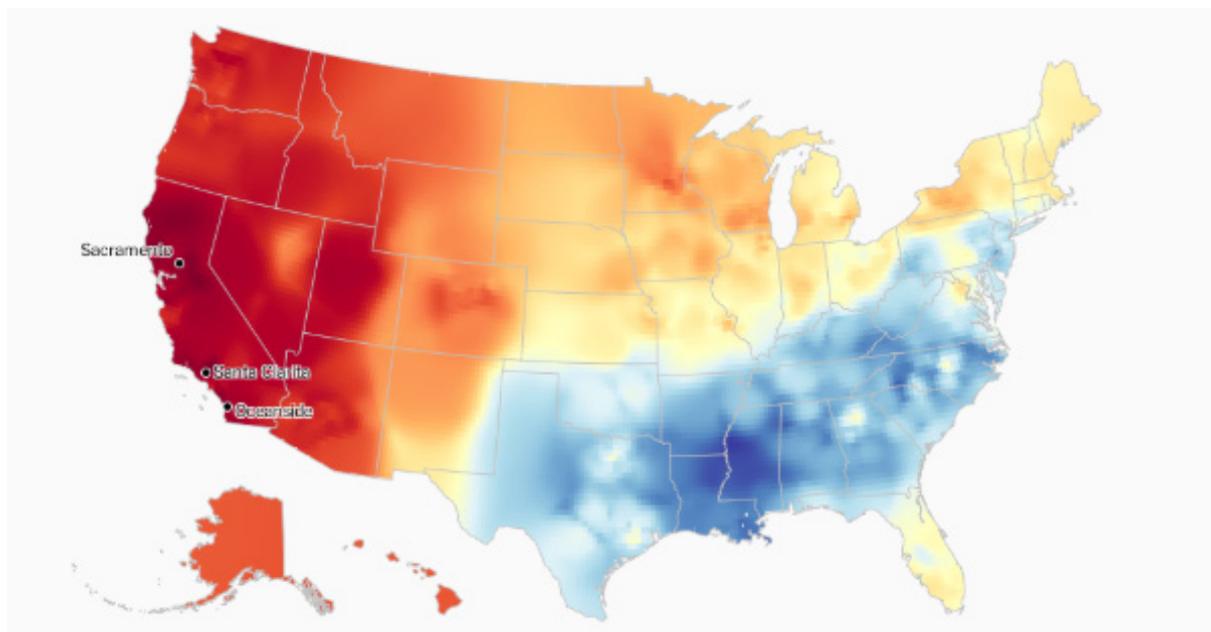
Fonte: *The Washington Post*

Observe que os designers desse material poderiam ter acrescentado um título descritivo seco (“Porcentagem de anos saudáveis versus expectativa de vida”) e plotar os dados em um gráfico de dispersão sem maiores esclarecimentos. Nesse caso, os leitores teriam que descobrir as informações por si mesmos e precisariam fazer um esforço para entender o real significado do gráfico.

Os designers desse projeto, porém, adotaram uma abordagem diferente: escolheram um título muito mais interessante e envolvente (“Quanto tempo vamos viver - e com que qualidade”) para conquistar e interessar os leitores. Além disso, também destacaram em pontos alguns países no próprio gráfico para adicionar contexto (“Estados Unidos: os homens podem esperar perder cerca de 10 anos para se tornarem incapazes”). Esta é a camada de anotação no trabalho.

Autores como Nick Diakopoulos (2013), da Northwestern University, sugerem orientações sobre como acrescentar/editar boas anotações. É o caso, por exemplo, o gráfico da Figura 8 que inclui muitos destaques textuais.

Figura 8: Gráfico sobre o stock Apple Stock



Fonte: *Apple Stock*

Conclusão

Qualquer pessoa que desenha gráficos, tabelas ou mapas nunca deve esquecer que a visualização vai bem mais além do que a visualização de dados para análise ou comunicação. Esse tipo de material deve contribuir para que os dados façam sentido para as pessoas, enfatizando o que é importante em relação ao que é secundário e, sempre que possível, colocando os dados em contexto e explicando o que significam. Como Ben Shneiderman afirma em entrevista concedida em 2019, “o propósito da visualização é a percepção, não olhar imagens” (8). Podemos apresentar as nossas percepções por meio de recursos visuais, palavras ou através de combinação de ambos. Se atingirmos esse objetivo nossos leitores certamente ficarão bastante agradecidos.

Referências

- CAIRO, Alberto. *The Truthful Art: Data, Charts, and Maps for Communication*. New Riders. Disponível em: <<https://www.amazon.es/Truthful-Art-Data-Charts-Communication/dp/0321934075>>. Acesso em: 14/08/2019.
- DATA@URBAN. *Três maneiras de anotar seus gráficos*. Medium Corporation. 17 de março de 2018. Disponível em: <https://medium.com/@urban_institute/three-ways-to-annotate-your-graphs-d140eo4e48ec>. Acesso em: 14/08/2019.
- DIAKOPOULOS, Nick. *Contaçõ de histórias com visualização de dados: o Contexto é o Rei*. Nick Diakopoulos. 17 de setembro de 2013. Disponível em: <<http://www.nickdiakopoulos.com/2013/09/17/storytelling-with-data-visualization-context-is->

king/>. Acesso em: 14/08/2019.

GRAFF, Ryan. *Por trás do mapa dialeto interativo: como um estagiário criou o conteúdo mais popular do New York Times em 2013*. Knight Lab. 20 de janeiro de 2014. Disponível em: <<https://knightlab.northwestern.edu/2014/01/20/behind-the-dialect-map-interactive-how-an-intern-created-the-new-york-times-most-popular-piece-of-content-in-2013/>>. Acesso em: 14/08/2019.

HULLMAN, Jessica. *O objetivo da visualização é visão, não imagens: Uma entrevista com o pioneiro da visualização Ben Shneiderman*. Medium Corporation. 12 de março de 2019. Disponível em: <<https://medium.com/multiple-views-visualization-research-explained/the-purpose-of-visualization-is-insight-not-pictures-an-interview-with-visualization-pioneer-ben-beb15b2d8e9b>>. Acesso em: 14/08/2019.

KATZ, Josh; ANDREWS, Wilson. *How Y'all, Youse and You Guys Talk*. *The New York Times*. 21 de dezembro de 2013. Disponível em: <<https://www.nytimes.com/interactive/2014/upshot/dialect-quiz-map.html>>. Acesso em: 14/08/2019.

WILKE, Claus O.; *Visualizando dados: Mapeando dados para estética*. Livro Fundamentos da visualização de dados. O'Reilly Media, Inc. Disponível em: <<https://serialmentor.com/dataviz/aesthetic-mapping.html>>. Acesso em: 14/08/2019.

Traduzido por Gustavo Pereira e Sonia Virgínia Moreira