

Importância, tendências e competências

em visualização de dados

Prof. Walmes Zeviani

walmes@ufpr.br

Laboratório de Estatística e Geoinformação
Departamento de Estatística
Universidade Federal do Paraná

10 de junho de 2019

Por que visualização de dados é tão importante?

O poder de ver além

Tabela 1. Os 4 pares de variáveis do quarteto de Ancombe.

x_1	y_1	x_2	y_2	x_3	y_3	x_4	y_4
10	8.04	10	9.14	10	7.46	8	6.58
8	6.95	8	8.14	8	6.77	8	5.76
13	7.58	13	8.74	13	12.74	8	7.71
9	8.81	9	8.77	9	7.11	8	8.84
11	8.33	11	9.26	11	7.81	8	8.47
14	9.96	14	8.10	14	8.84	8	7.04
6	7.24	6	6.13	6	6.08	8	5.25
4	4.26	4	3.10	4	5.39	19	12.50
12	10.84	12	9.13	12	8.15	8	5.56
7	4.82	7	7.26	7	6.42	8	7.91
5	5.68	5	4.74	5	5.73	8	6.89

O poder de ver além

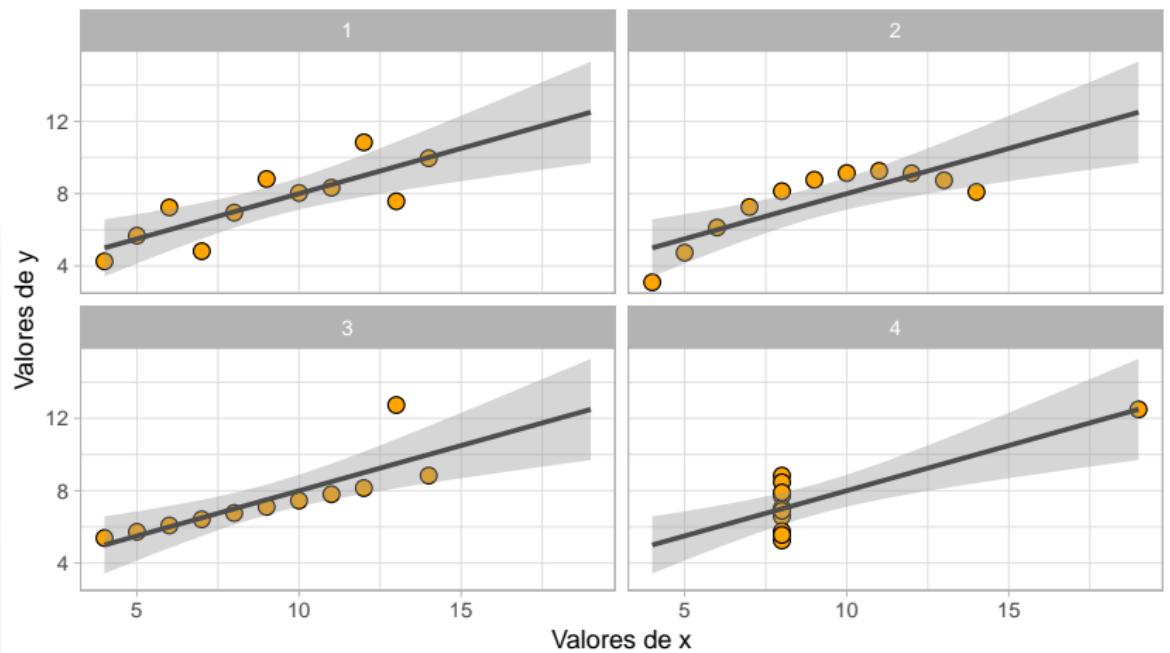


Figura 1. O quarteto de Anscombe em um diagrama de dispersão.

Rápida percepção, rápida reação

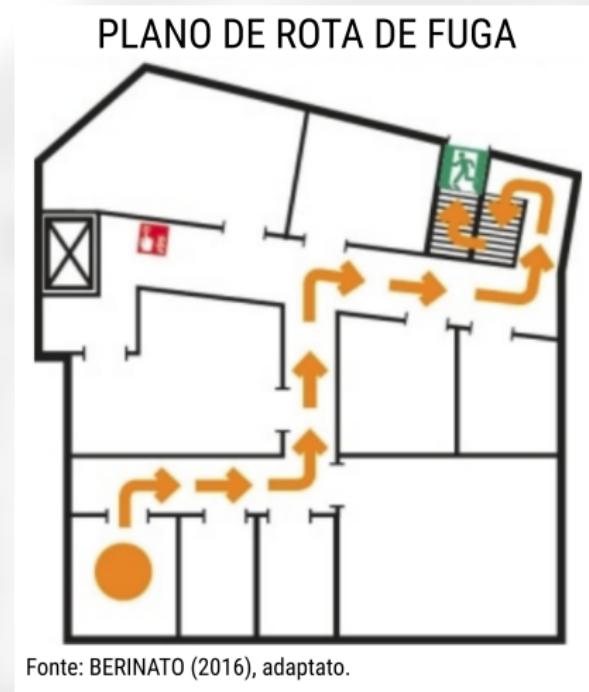


Figura 2. Mapa de rota de fuga. Adapatado de Berinato (2016).

Rápida percepção, rápida reação

Saia desta sala. Vire à direita e caminhe 3 metros até o final do corredor, onde você estará em frente a uma grande sala de conferências. Vire à esquerda e caminhe mais 3 metros até chegar ao final do corredor. À sua esquerda estará um alarme de incêndio, perto do elevador. À sua direita, no final do corredor, há uma escadaria. Não vá ao elevador. Vire à direita e caminhe mais 4 metros até o final do corredor, vire à esquerda e desça pelas escadas. Desça dois lances de escadas e saia do prédio na porta na parte inferior das escadas.

Figura 3. Plano de rota de fuga. Adapatado de Berinato (2016).



*“O maior valor de um gráfico é quando ele nos força a **perceber** o que nós jamais esperávamos ver.”*

– John Tukey · Estatístico



*“Visualizações funcionam como uma fogueira na qual nos reunimos para **contar histórias**.”*

– Al Shalloway · CEO Net Objectives



*“A maioria de nós precisa ouvir a música para **entender** o quão bonita ela é. No entanto, é assim que apresentamos estatísticas: nós apenas mostramos as notas, nós não tocamos a **música**.”*

– Hans Rosling

O fenômeno Big Data

- ▶ **Big data:** conjuntos de dados grandes ou complexos na qual as abordagens tradicionais de processamento são inadequadas/impraticáveis.
- ▶ **Desafios:** captura, armazenamento, qualidade, beneficiamento, consulta, transferência, **visualização**, análise, privacidade e governança.



Figura 4. Os 8 V's do Big Data.

Visualização de dados

Definição de como deve ser



Martin LeBlanc
@martinleblanc

Seguir



data visualization

A ~~user interface~~ is like a joke. If you have to explain it, it's not that good.

10:56 - 14 de mai de 2014

2.811 Retweets 1.939 Curtidas



61

2,8 mil

1,9 mil

Definição

Visualização de dados

A **representação** e **apresentação** de dados que exploram nossa habilidade de **percepção visual** com a finalidade de aumentar a **compreensão** (KIRK, 2012).

Tipos de visualização

Segundo Berinato (2016), dois aspectos determinam o tipo de visualização:

1. A informação é **conceitual** ou **orientada a dados**?
2. Está se **declarando** ou **explorando** algo?

Os 4 tipos de visualização



Figura 5. Os 4 tipos de visualização que resultam das respostas às duas perguntas.
Fonte: adaptado de Berinato (2016).

Desenvolvimento de ideia

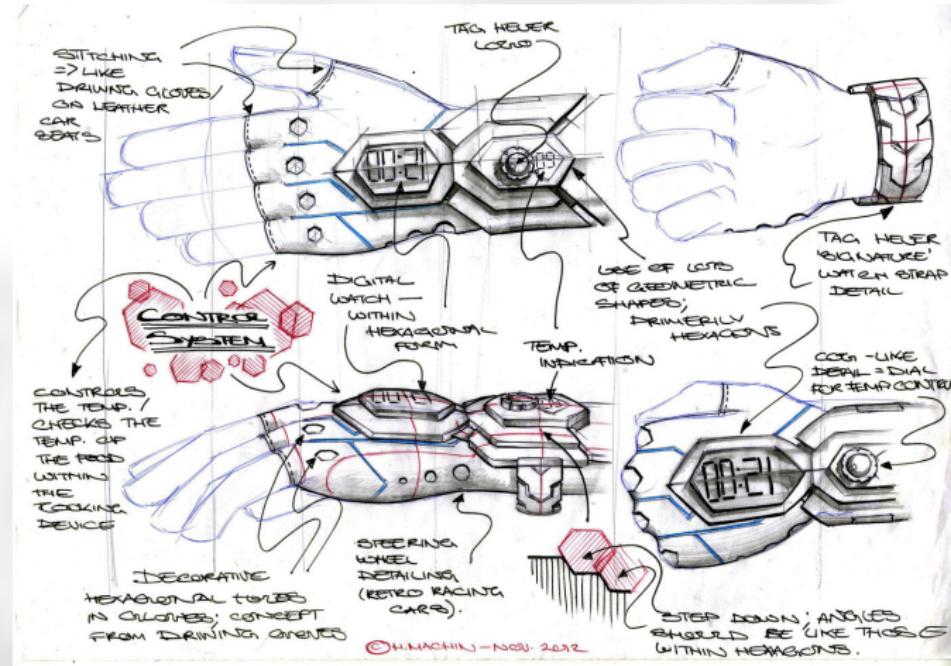


Figura 6. Fonte:

<<https://www.coroflot.com/harriettemachin/TAG-Heuer-Branding-Project>>.

Ilustração de ideia

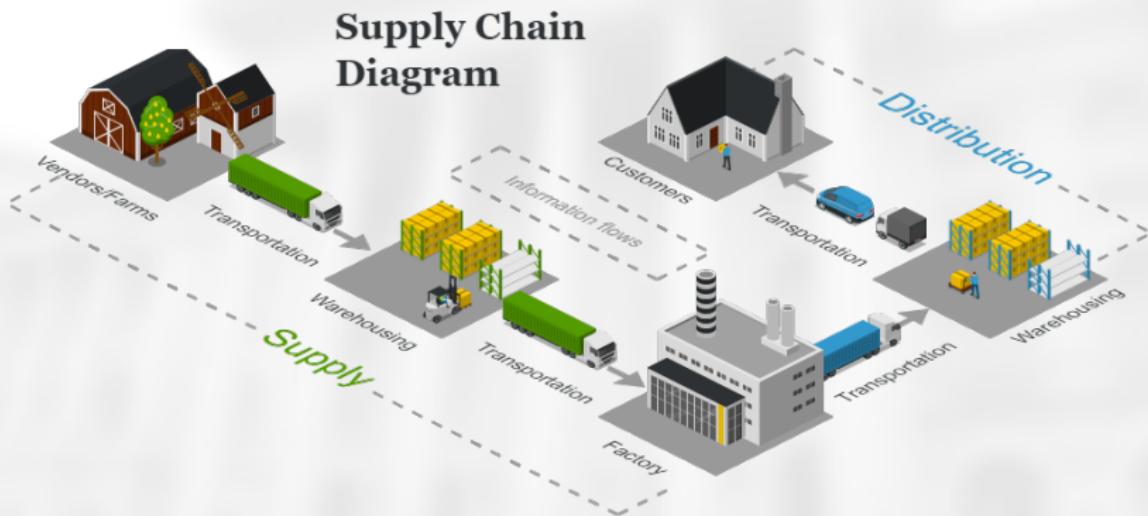


Figura 7. Fonte: <<https://icograms.com/usage-supply-chain-diagram.php>>.

Descoberta visual

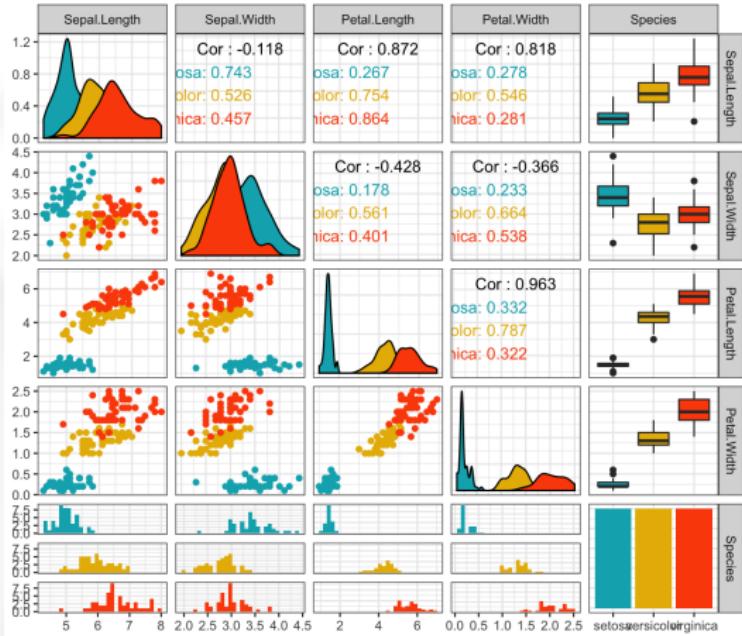


Figura 8. Fonte: <<http://www.sthda.com/english/articles/32-r-graphics-essentials/130-plot-multivariate-continuous-data/>>.

Visualização do cotidiano

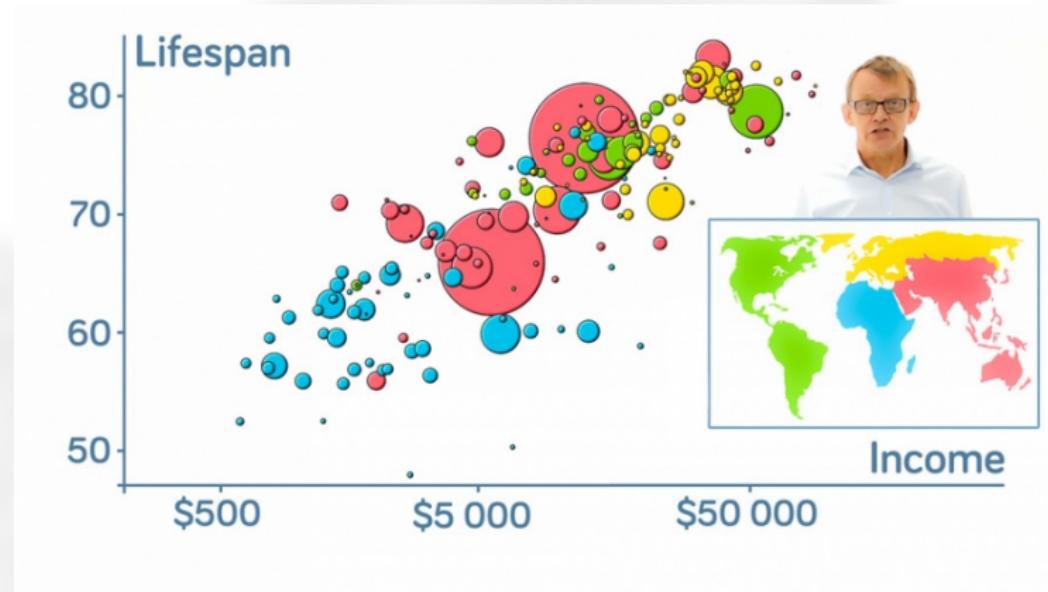


Figura 9. Fonte: <<http://inthesetimes.com/article/21771/new-optimists-bill-gates-steven-pinker-hans-rosling-world-health>>.

A visualização orientadas a dados

A **escolha de um gráfico** depende de uma série de elementos. Os principais são:

- ▶ Escala das variáveis.
- ▶ Propósito da visualização → função do gráfico.
- ▶ Questões visuais/estéticas → forma do gráfico.

Os 4 tipos de escala

Qualitativa

- Níveis categóricos.

Quantitativa

- Níveis métricos.

Nominal

- Sem ordenação.
- Sexo, nacionalidade, tipo de sangue, religião.

Ordinal

- Com ordenação.
- Escala de Likert, satisfação, cargo/patente.

Discreta

- Suporte nos números inteiros.
- Contagens.

Contínua

- Suporte nos números reais.

Informação

Intervalar

- Diferenças têm significado.
- Temperatura °C e °F, data/hora, latitude/longitude.

Razão

- Diferenças e razões têm significado.
- Temperatura Kelvin, peso, distância, renda, duração.

Forma e função de um gráfico

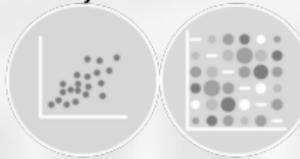
- ▶ Agrupamento dos gráficos baseado nas características comuns.
- ▶ **Função**
 - ▶ Classificação principal: função do gráfico.
 - ▶ Relacionado às quantidades a serem representadas.
 - ▶ Quais quantidades/propriedades serão exibidas → **o que?**
- ▶ **Forma**
 - ▶ Classificação secundária: forma do gráfico.
 - ▶ Relacionado às decisões estéticas e de mídia.
 - ▶ Como as quantidades/propriedades serão exibidas → **como?**

Função de um gráfico

Distribuição



Relação



Ordenação



Evolução



Mapa



Organização



Fluxo



Função de um gráfico

Distribuição Caracterizam a **distribuição ou divisão** de variáveis qualitativas ou quantitativas.

Relação Exibem a forma da **relação ou contraste** entre variáveis.

Ordenação Priorizam exibir a **classificação** de categorias/elementos baseada nas variáveis.

Evolução Exploram o comportamento de dados **cronológicos**.

Mapa Representam dados com propriedades **geográficas**.

Organização Retratam a **estrutura** de organização de coleções de categorias/elementos.

Fluxo Descrevem o **movimento ou conexão** entre categorias/elementos.

Variáveis visuais

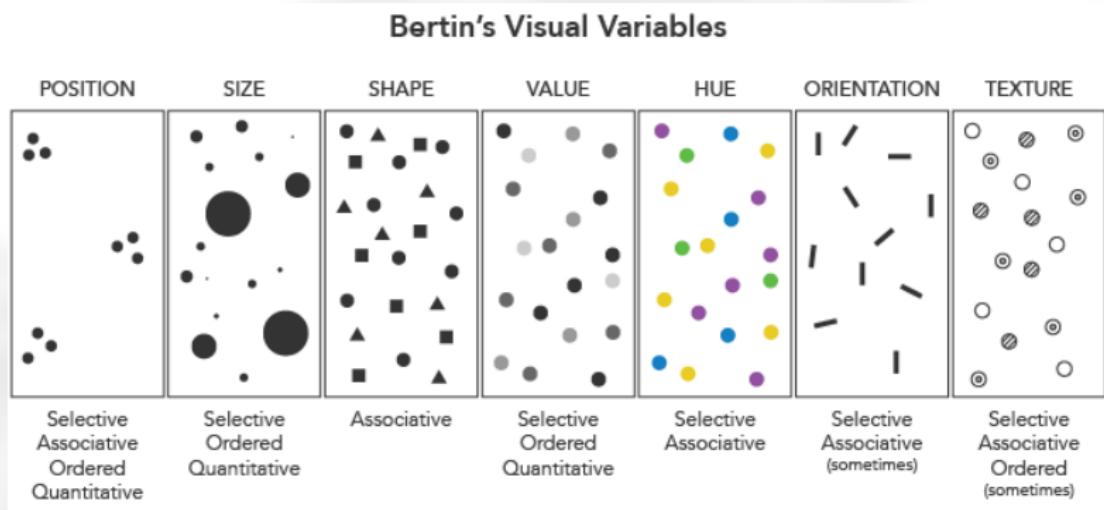


Figura 10. As 7 variáveis visuais definidas por Bertin.

As **7 variáveis visuais**: posição, tamanho, forma, saturação, cor, orientação e textura.¹

¹www.axismaps.com

Propriedades das variáveis visuais

1. Seletiva.
2. Associativa.
3. Quantitativa.
4. Ordenativa.
5. Dissociativa.

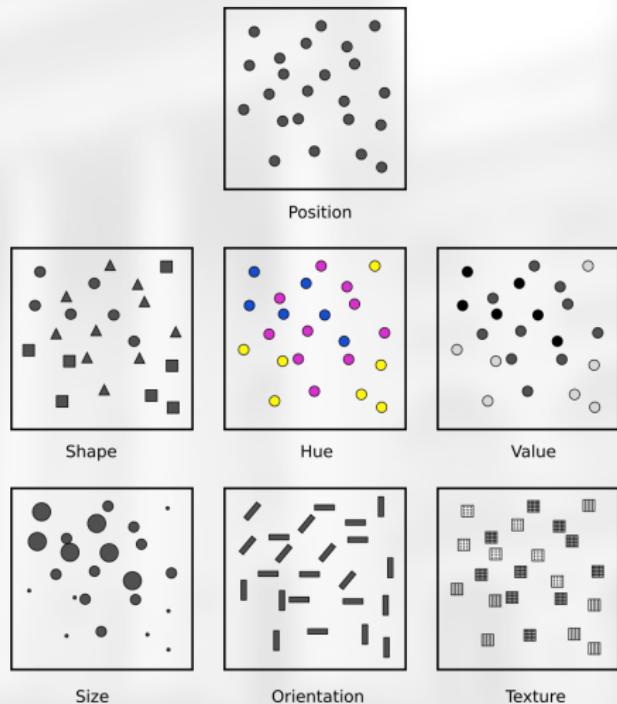


Figura 11. As 7 variáveis visuais.

Recomendação de uso

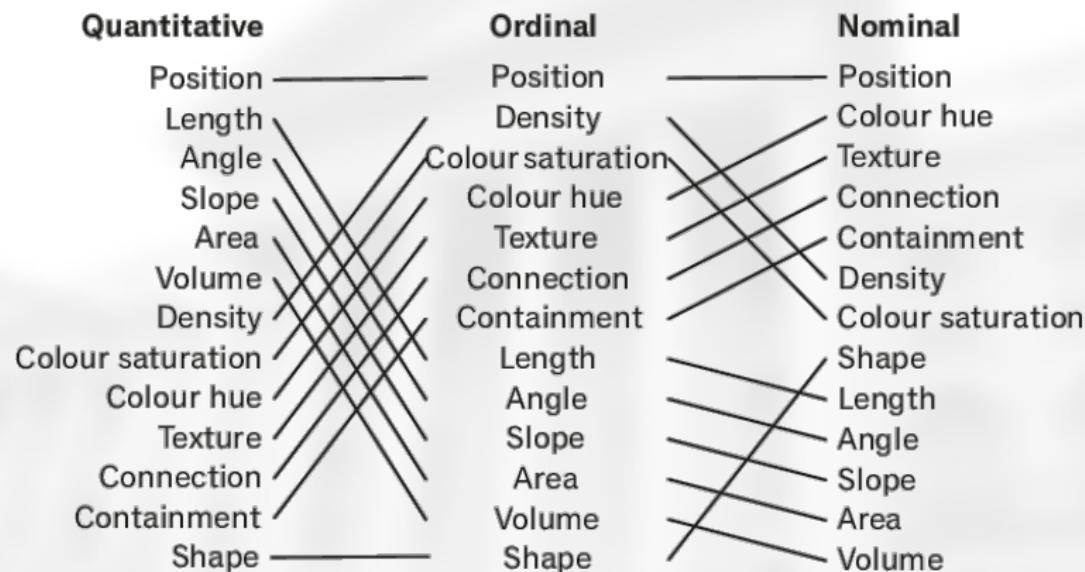


Figura 12. Ranking de Mackinlay das variáveis visuais para o mapeamento de variáveis quantitativas, qualitativas ordinais e nominais.

<https://www.researchgate.net/publication/312530613_Images_of_Time>

Tendências em visualização de dados

Linha do tempo da Visualização de Dados

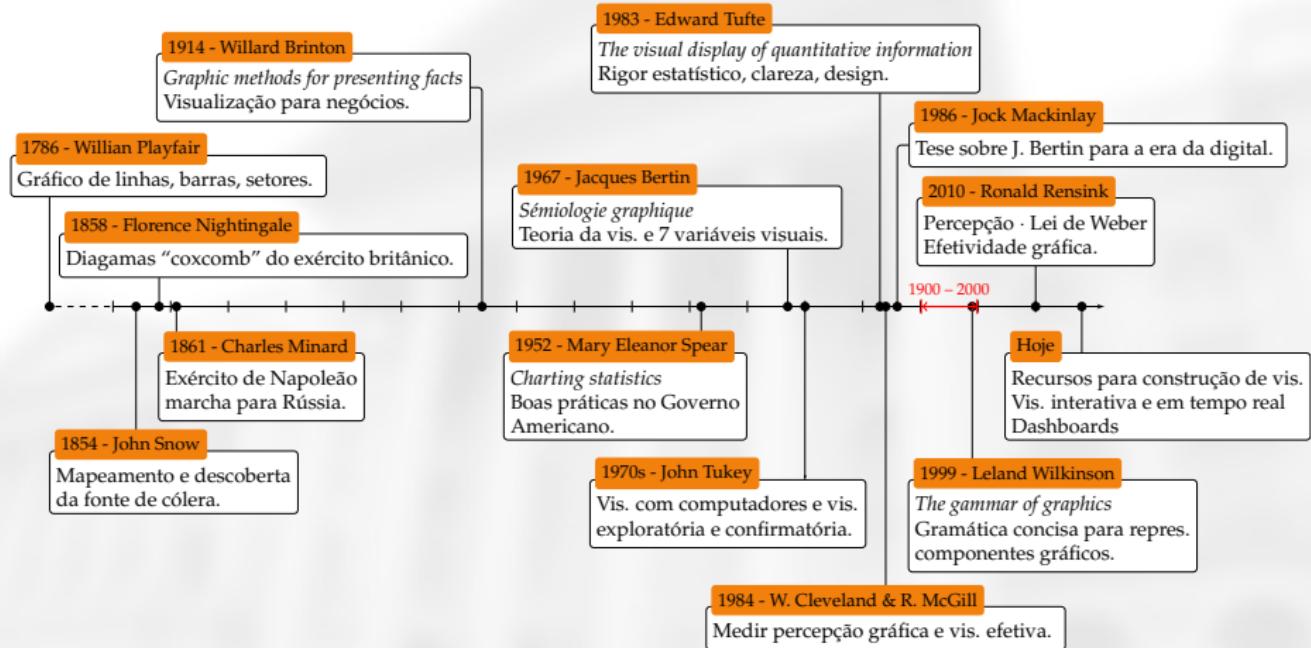


Figura 13. A linha do tempo da visualização de dados. Fonte: o autor.

O presente

- ▶ 2010s
 - ▶ softwares acessíveis + disposição de dados → democratiza a prática da visualização.
 - ▶ Cria-se uma atmosfera propícia para experiências.
 - ▶ Visualização não é mais uma província de experts, é um fenômeno.
- ▶ Hoje
 - ▶ Recursos/ferramentas evoluem.
 - ▶ Representações gráficas mais apropriadas são desenvolvidas.
 - ▶ Novas dimensões: animação, reatividade e interatividade.
 - ▶ Velocidade e acesso:
 - ▶ Atualização em tempo real.
 - ▶ Relatórios web dinâmicos.
 - ▶ Smart things: smartwatch, apps esportivos.
 - ▶ Monitoramento com dashboards.

Paineis de visualização · Dashboards

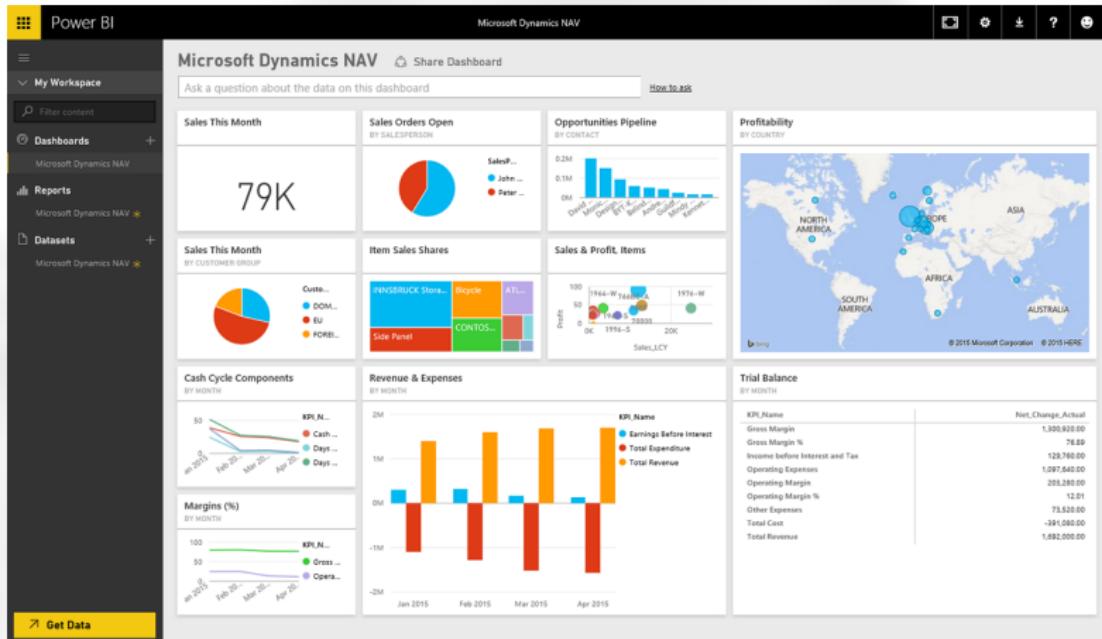


Figura 14. Exemplo de dashboard produzido com o software Power BI.

Paineis de visualização · Dashboards

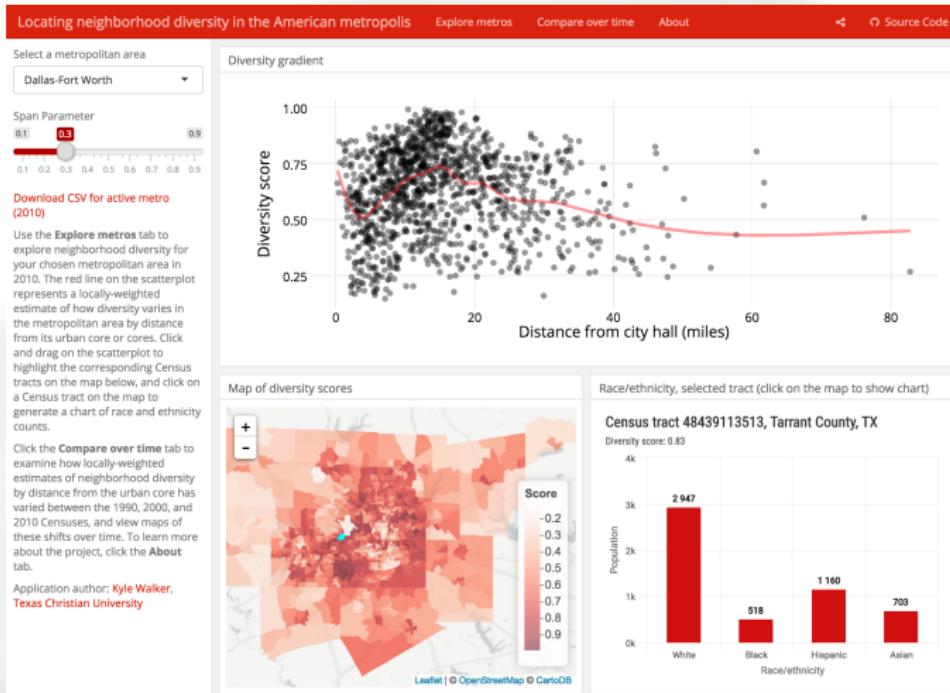


Figura 15. Exemplo de dashboard feito com o R usando o pacote shiny.

Competências em visualização de dados

Os 8 chapéus da visualização de dados

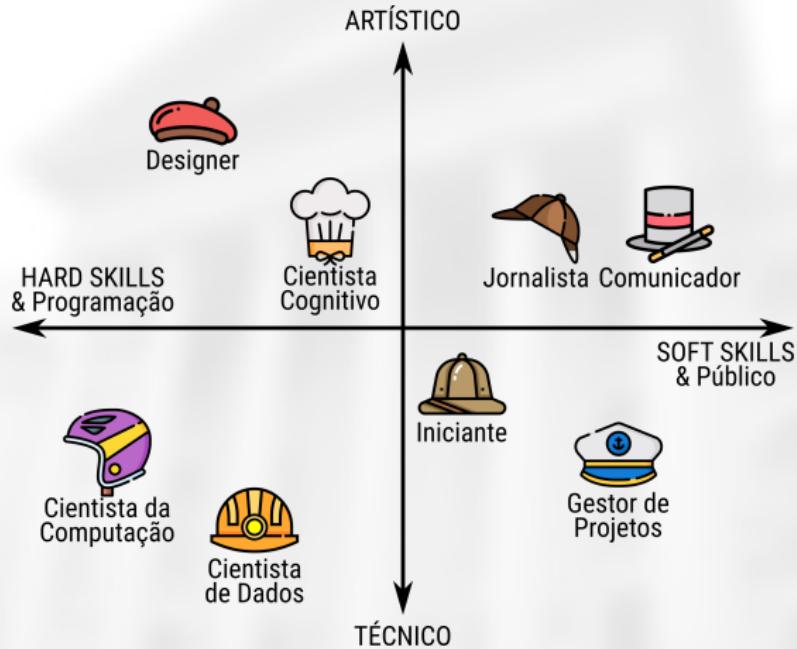


Figura 16. Os 8 “chapeús” de Kirk considerando os eixos: competência técnica-artística sobre visualização de dados e habilidades de programação (hard skills) e comunicação/gestão (soft skills). Fonte: o autor.

Considerações finais

Recomendações

“Os números tem uma importante história para contar. Eles dependem de você dar-lhes um clara e convincente voz.”

– Stephen Few

- ▶ Esforce-se para compreender melhor os seus dados: veja-os.
- ▶ Uma apropriada visualização impacta mais do que múltiplas tabelas, testes e índices de ajuste.
- ▶ Invista em recursos para **manipulação e visualização** de dados: → **R**, Python, Inkscape, Gimp, JavaScript.

Referências

-  BERINATO, S. **Good Charts: The HBR Guide to Making Smarter, More Persuasive Data Visualizations.** Harvard Business Review Press, 2016. ISBN 9781633690714. Disponível em: <https://books.google.com.br/books?id=_kJBCgAAQBAJ>.
-  KIRK, A. **Data Visualization: A Successful Design Process.** Packt Pub., 2012. (Community experience distilled). ISBN 9781849693479. Disponível em: <<https://books.google.com.br/books?id=l4qBVLfD3t4C>>.

Pela oportunidade e atenção, agradeço.

Prof. Walmes Zeviani
<walmes@ufpr.br>
<leg.ufpr.br/~walmes>