

Introdução à Visualização

SCC0652 – Visualização Computacional

Profa. Maria Cristina
cristina@icmc.usp.br

Instituto de Ciências Matemáticas e de Computação (ICMC)
Universidade de São Paulo (USP)

VICG Grupo de Visualização,
ICMC USP Imagens e Computação Gráfica

Sumário

1 Motivação

2 Panorama

- O que é Visualização (de dados)
- Breve histórico
- Visualização nos dias atuais
- Relação com outras disciplinas
- O *pipeline* de Visualização
- O papel da percepção em Visualização

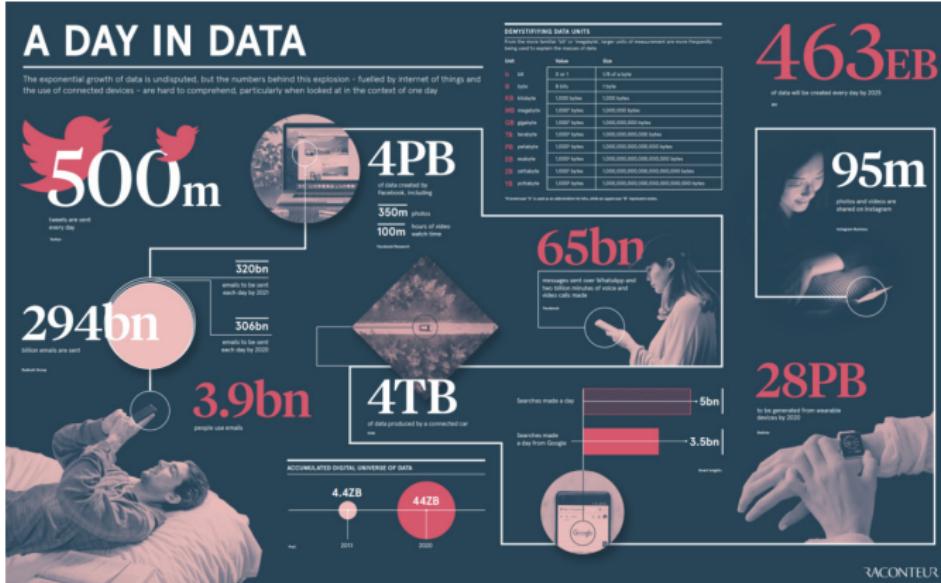
3 Objetivos de uma visualização

4 Referências

Motivação

- Volume de dados digitais é impressionante, e crescendo mais rápido do que nunca!
- Artigo recente da Forbes afirmava que até o ano de 2020, seriam criados cerca de 1,7 megabytes de novos dados **a cada segundo para cada ser humano no planeta**

Motivação



- Volume de dados digitais é impressionante!
- (Abril 2019) <https://www.visualcapitalist.com/how-much-data-is-generated-each-day/>

Motivação

Alguns números no infográfico:

- 500 milhões de tweets são enviados
- 294 bilhões de emails são enviados
- 4 petabytes de dados são criados no Facebook
- 4 terabytes de dados são gerados em cada carro conectado
- 65 bilhões de mensagens são enviadas no WhatsApp
- 5 bilhões de buscas são feitas

Outra estimativa afirma que em 2025 serão criados diariamente no globo 463 exabytes de dados, o que equivale a 212,765,957 DVDs por dia!

E se considermos o que acontece a cada minuto...

Motivação



[https://fossbytes.com/
how-much-data-is-generated-every-minute-in-the-world/](https://fossbytes.com/how-much-data-is-generated-every-minute-in-the-world/)

Motivação

Abbreviation	Unit	Value	Size (in bytes)
b	bit	0 or 1	1/8 of a byte
B	bytes	8 bits	1 byte
KB	kilobytes	1,000 bytes	1,000 bytes
MB	megabyte	$1,000^2$ bytes	1,000,000 bytes
GB	gigabyte	$1,000^3$ bytes	1,000,000,000 bytes
TB	terabyte	$1,000^4$ bytes	1,000,000,000,000 bytes
PB	petabyte	$1,000^5$ bytes	1,000,000,000,000,000 bytes
EB	exabyte	$1,000^6$ bytes	1,000,000,000,000,000,000 bytes
ZB	zettabyte	$1,000^7$ bytes	1,000,000,000,000,000,000,000 bytes
YB	yottabyte	$1,000^8$ bytes	1,000,000,000,000,000,000,000,000 bytes

o Desafio...

Como fazer (bom) uso de tanto dado?

Algumas 'buzzwords'

- *data science*
- *big data*
- *data analytics*

Algumas 'buzzwords'

- *data science*
 - *big data*
 - *data analytics*
-
- O que significam?

Algumas 'buzzwords'

- *data science*
- *big data*
- *data analytics*
- Leitura: [https://www.simplilearn.com/
data-science-vs-big-data-vs-data-analytics-article](https://www.simplilearn.com/data-science-vs-big-data-vs-data-analytics-article)

Responda

- o que você entende por *data science*
- o que você entende por *big data*
- o que você entende por *data analytics*

Sumário

1 Motivação

2 Panorama

- O que é Visualização (de dados)
- Breve histórico
- Visualização nos dias atuais
- Relação com outras disciplinas
- O *pipeline* de Visualização
- O *papel da percepção* em Visualização

3 Objetivos de uma visualização

4 Referências

Sumário

1 Motivação

2 Panorama

- O que é Visualização (de dados)
- Breve histórico
- Visualização nos dias atuais
- Relação com outras disciplinas
- O *pipeline* de Visualização
- O papel da percepção em Visualização

3 Objetivos de uma visualização

4 Referências

O que é Visualização

“Visualização é a comunicação de informação por meio de representações gráficas” [Ward et al., 2010]

O que é Visualização

“Visualização é a comunicação de informação por meio de representações gráficas” [Ward et al., 2010]

- Uma **única imagem** pode embutir **grande quantidade de informação** e ser interpretada muito mais rapidamente do que um texto
 - **Interpretação de imagens** é realizada **em paralelo** pelo sistema perceptual do cérebro humano, enquanto a leitura de texto é sequencial (leitura)
 - Uma imagem também é menos dependente da língua: forma de comunicação mais universal

O que é Visualização

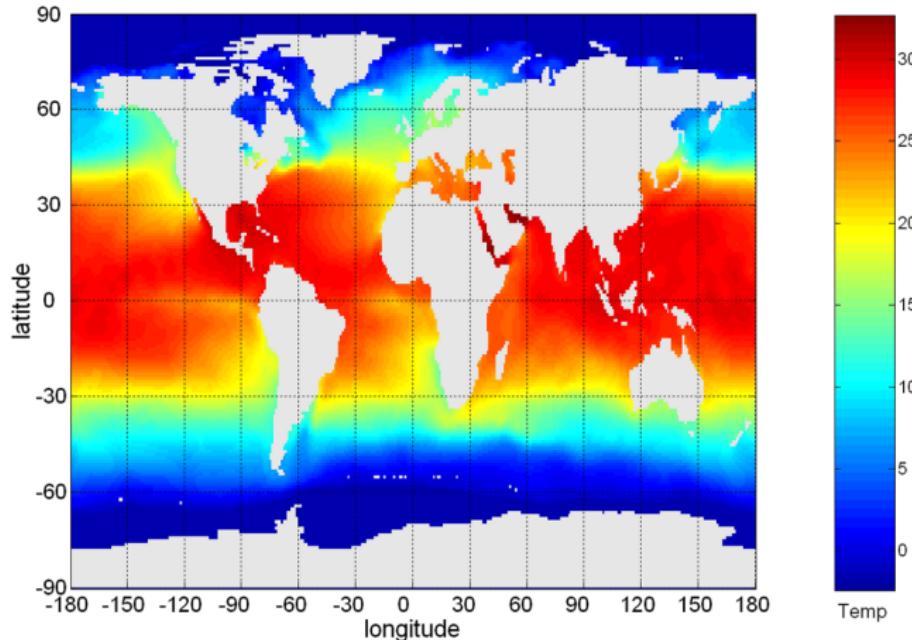


Figura: 10.000 medidas de temperatura na superfície do oceano resumidas em uma única figura.

Visualização no dia-a-dia

Representações gráficas são muito empregadas em diversas situações em **substituição** à divulgação de **informação** de forma **verbal ou escrita**

Visualização no dia-a-dia

Representações gráficas são muito empregadas em diversas situações em **substituição** à divulgação de **informação** de forma **verbal ou escrita**

- Atividades cotidianas
 - Mapas de trem e metrô
 - Mapa de uma região para determinar rota
 - Gráficos explicativos em jornais e revistas (infográficos)
 - Gráficos de previsão do tempo
 - Imagens de tomógrafos computadorizados
 - Manuais de instrução para montagem de móveis, bicicletas, etc.
- Atividades industriais/profissionais
 - Análise do mercado de ações
 - Desenhos de engenharia mecânica e civil
 - Diagnóstico por imagem, p.ex. de câncer de mama
 - Simulação de processos complexos

Por que Visualizar?

- Apresentar estatísticas dos dados não seria suficiente?
- A maneira de apresentar os dados influencia um processo de tomada de decisão?
 - Pode modificar uma decisão?
 - Pode induzir decisões erradas?
 - Existe alguma representação melhor?

Por que Visualizar?

Anscombe's Quartet: Raw Data

	1		2		3		4	
	X	Y	X	Y	X	Y	X	Y
	10.0	8.04	10.0	9.14	10.0	7.46	8.0	6.58
	8.0	6.95	8.0	8.14	8.0	6.77	8.0	5.76
	13.0	7.58	13.0	8.74	13.0	12.74	8.0	7.71
	9.0	8.81	9.0	8.77	9.0	7.11	8.0	8.84
	11.0	8.33	11.0	9.26	11.0	7.81	8.0	8.47
	14.0	9.96	14.0	8.10	14.0	8.84	8.0	7.04
	6.0	7.24	6.0	6.13	6.0	6.08	8.0	5.25
	4.0	4.26	4.0	3.10	4.0	5.39	19.0	12.50
	12.0	10.84	12.0	9.13	12.0	8.15	8.0	5.56
	7.0	4.82	7.0	7.26	7.0	6.42	8.0	7.91
	5.0	5.68	5.0	4.74	5.0	5.73	8.0	6.89
Mean	9.0	7.5	9.0	7.5	9.0	7.5	9.0	7.5
Variance	10.0	3.75	10.0	3.75	10.0	3.75	10.0	3.75
Correlation	0.816		0.816		0.816		0.816	

Figura: Fonte: T. Munzner Visualization Analysis & Design (Fig. 1.3).

Por que Visualizar?

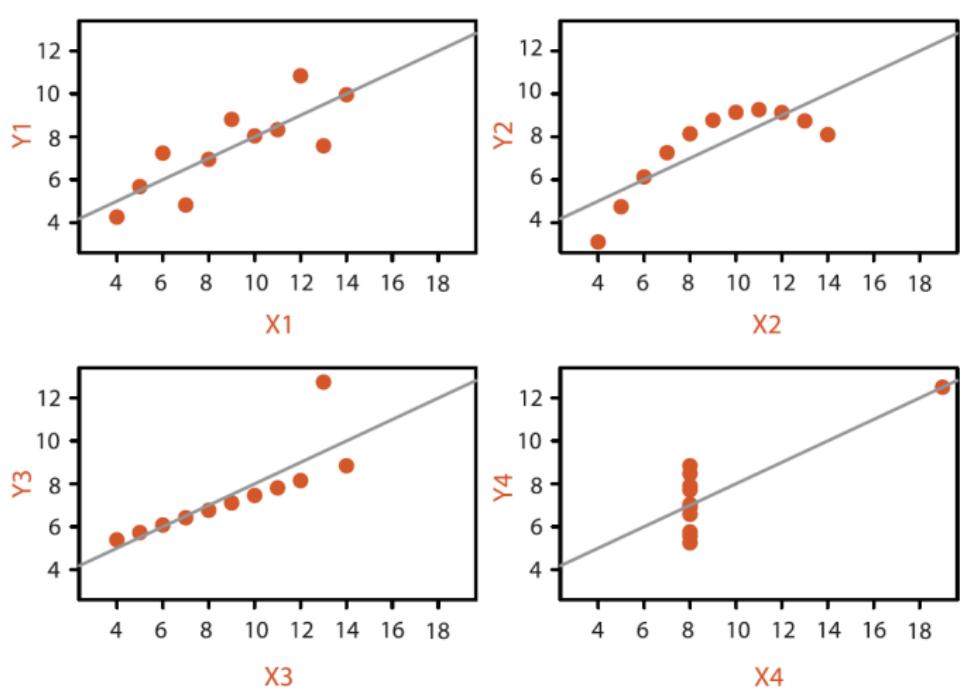


Figura: Fonte: T. Munzner Visualization Analysis & Design (Fig. 1.3).

Por que Visualizar?

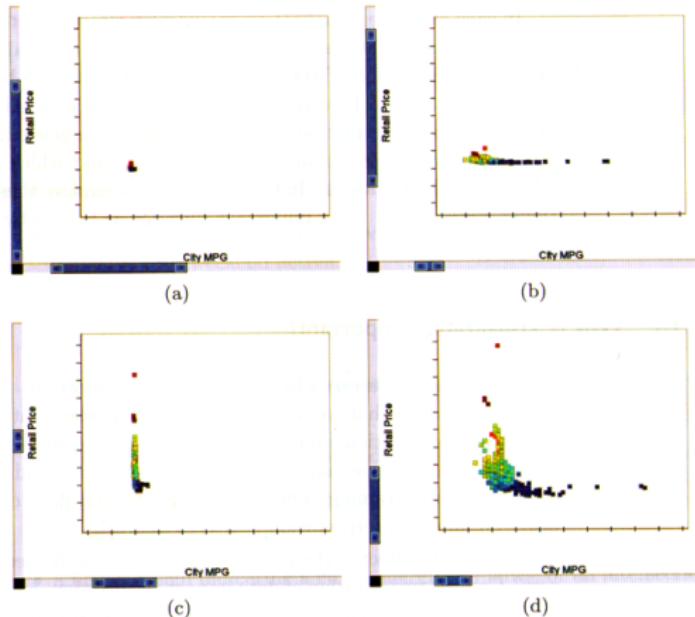


Figura: Gráficos de dispersão exibindo os mesmos dados (sobre carros), mas considerando diferentes intervalos de variação (*ranges*) nos eixos em *x* e *y*. (a) intervalos muito grandes em *x* e *y*. (b) intervalo grande em *y*, adequado em *x*. (c) intervalo muito grande em *x*, adequado em *y*. (d) intervalos adequados em *x* e em *y*.

Por que Visualizar

- A apresentação gráfica dos dados pode revelar muito mais do que os números em si!
- A maneira de apresentar impacta a percepção dos dados e o processo de análise!
 - Já demonstrado em estudos com usuários

Por que Visualizar

- A apresentação gráfica dos dados pode revelar muito mais do que os números em si!
- A maneira de apresentar impacta a percepção dos dados e o processo de análise!
 - Já demonstrado em estudos com usuários

Estudo de Caso: comparação de tratamentos clínicos

- Comparação de dois tratamentos clínicos para uma mesma condição (convencional/em teste)
- Supostamente um deles é muito superior ao outro
 - Resultados comparativos dos tratamentos reportados de diferentes maneiras
 - Decisão a ser tomada: manter ou interromper os testes?

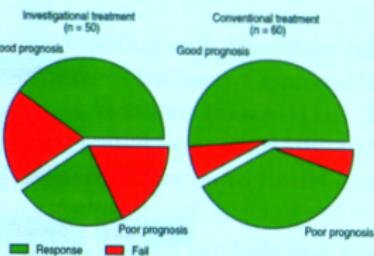
Por que Visualizar?

Table

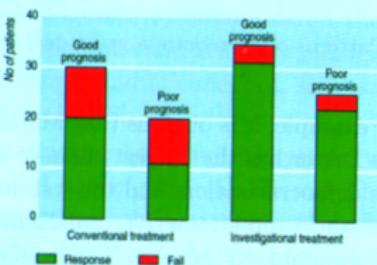
	Conventional treatment		Investigational treatment	
	Total no	% Fail	Total no	% Fail
Good prognosis	30	30	35	11
Poor prognosis	20	45	25	12
Total	50	38	60	12

(Negatively framed tables displayed failure rates in red.
Positively framed tables displayed response rates in green)

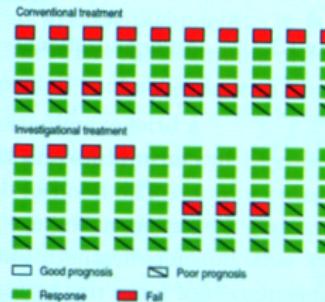
Pie chart



Bar graph



Icon



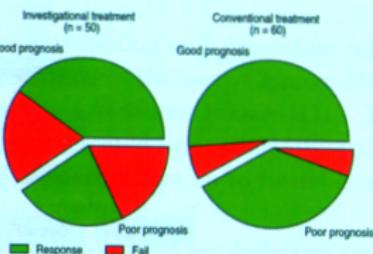
Por que Visualizar?

Table

	Conventional treatment		Investigational treatment	
	Total no	% Fail	Total no	% Fail
Good prognosis	30	30	35	11
Poor prognosis	20	45	25	12
Total	50	38	60	12

(Negatively framed tables displayed failure rates in red.
Positively framed tables displayed response rates in green)

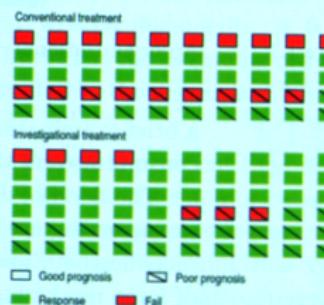
Pie chart



Bar graph



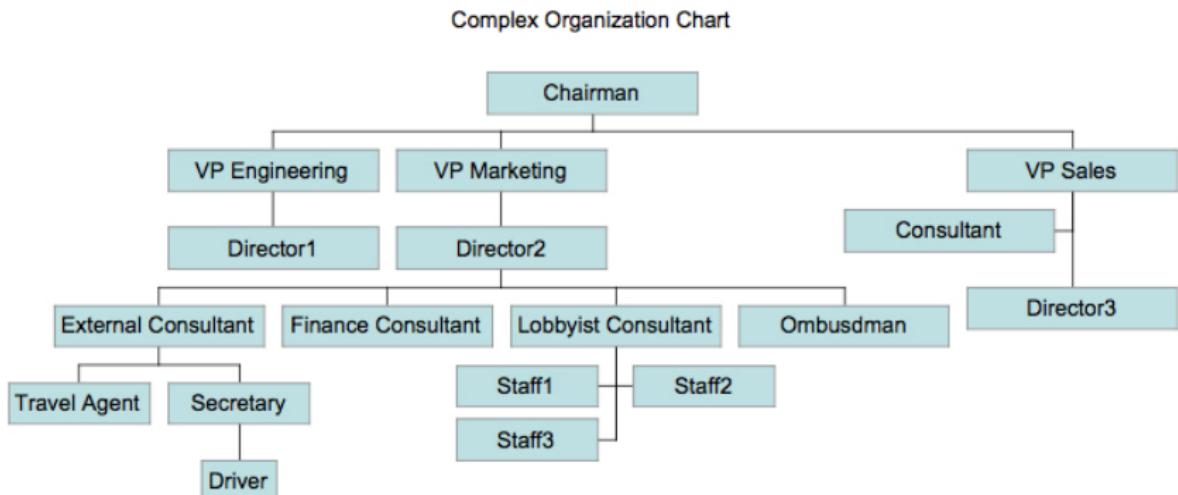
Icon



A representação com ícones (inferior direita) mostrou-se mais efetiva para a tomada da decisão (82% de acertos), e os gráficos de barra e pizza os menos efetivos (56% de acertos).

Por que Visualizar?

Visualização pode facilmente expressar certo tipo de informação que é difícil apresentar verbalmente



Por que Visualizar?

- A **importância** está na capacidade de **interpretar** dados mais **rapidamente**
- O canal visual humano é muito eficiente para captar informação relevante
- Auxílio essencial para apoiar processos de descoberta de **conhecimento** e **tomada de decisão**
- Mais do que infográficos: **interação** do analista!

Sumário

1 Motivação

2 Panorama

- O que é Visualização (de dados)
- **Breve histórico**
- Visualização nos dias atuais
- Relação com outras disciplinas
- O *pipeline* de Visualização
- O papel da percepção em Visualização

3 Objetivos de uma visualização

4 Referências

Primeiras Visualizações

Pinturas em cavernas datam de mais de 30,000 anos



Figura: Pintura em caverna às margens do rio Vézère, Pireneus, França.

Primeiras visualizações

Figuras já eram usadas para codificar palavras na antiguidade



Figura: Kish limestone tablet (Mesopotâmia).

Primeiras visualizações

Mapas de estradas do império Romano, com as distâncias
aproximadas e pontos de interesse

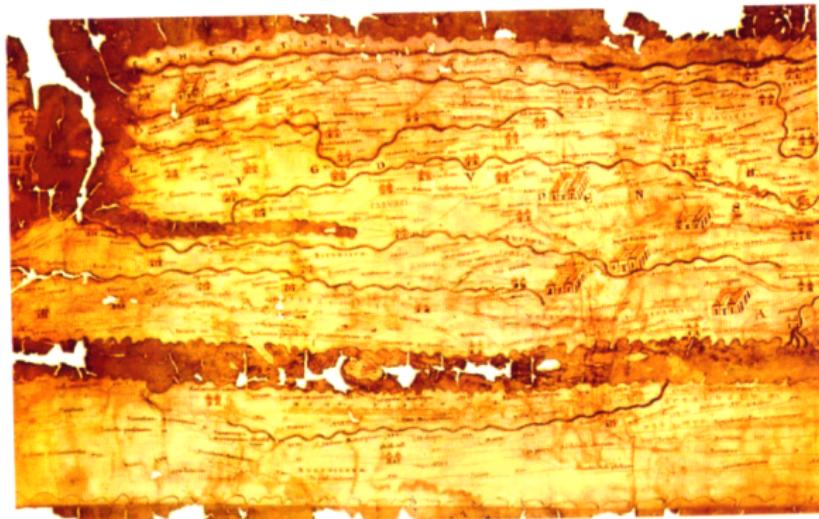


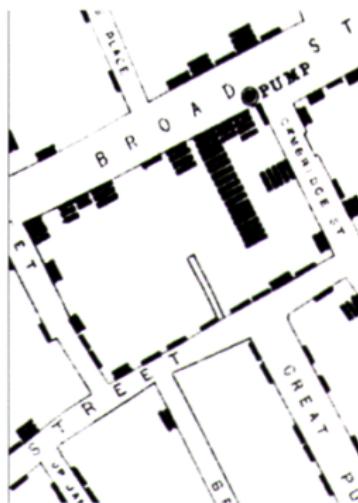
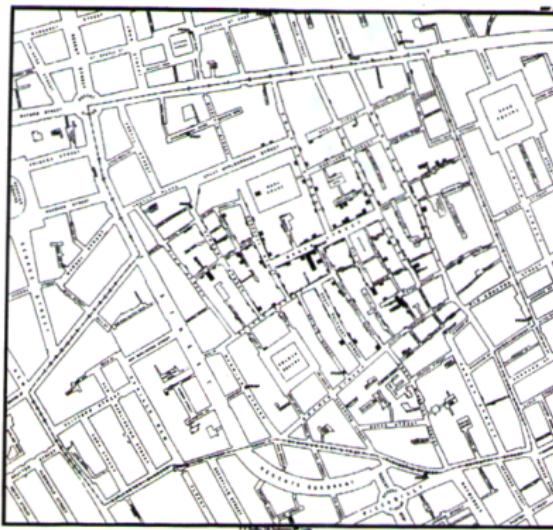
Figura: Peutinger Map (uma das 12 páginas).

Primeiras visualizações

- Mapas do mundo já eram desenhados na Idade Média
- Hereford Map, catedral de Hereford, País de Gales
- Data estimada entre 1276 e 1285, o maior mapa medieval conhecido
 - <https://www.themappamundi.co.uk/mappa-mundi/>
 - Jerusalém no centro do mundo

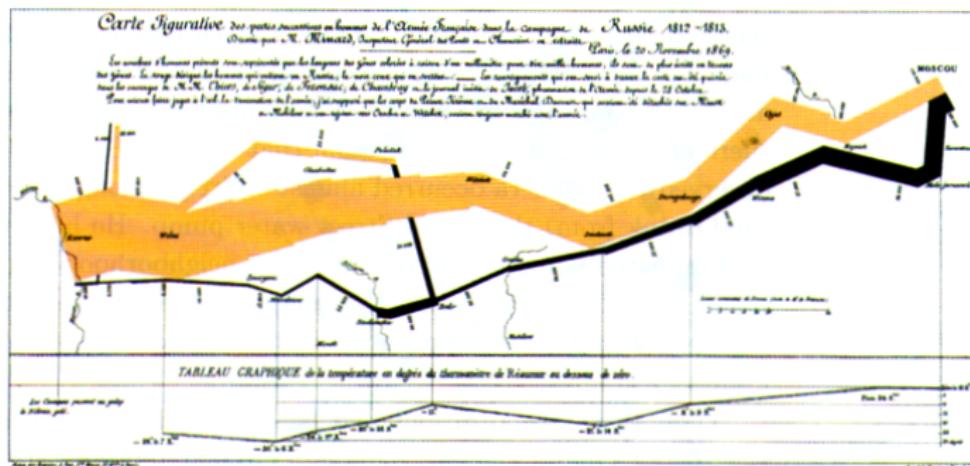
Primeiras visualizações

- Mapa de John Snow detalhando as mortes por cólera em Londres (em 1854)
 - Cada barra indica uma morte no local
 - Mais de 500 mortes verificadas na região de um poço na *Broad Street*



Primeiras visualizações

- Mapa de Minard sobre a expedição do exército de Napoleão na Rússia: de 400.000 soldados, somente 10.000 retornaram
- [https://scienzenorway.no/
blog-blog-from-numbers-to-graphics-statistics/
charles-joseph-minards-map-of-napoleons-flawed-russian-campaign-a-1618695](https://scienzenorway.no/blog-blog-from-numbers-to-graphics-statistics/charles-joseph-minards-map-of-napoleons-flawed-russian-campaign-a-1618695)



Primeiras visualizações

- Marco importante (William Playfair): introdução do conceito de eixos em mapas de gráficos de barra e séries temporais
- Representações visuais não mais restritas a mapeamentos geoespaciais...
- https://en.wikipedia.org/wiki/William_Playfair

Primeiras visualizações

Representação taxa de mortalidade mensal no Exército americano

- Em azul, mortes por doenças, em vermelho, mortes por ferimentos (batalha), em preto mortes por outras causas
- Veja:
<https://www.york.ac.uk/depts/math/histstat/small.htm>

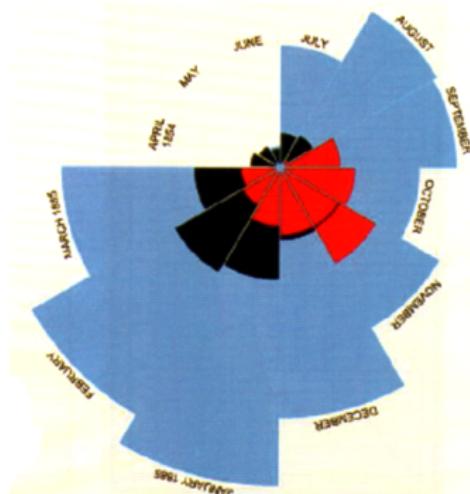


Figura: Coxcomb chart, Florence Nightingale. Circa 1858.

Sumário

1 Motivação

2 Panorama

- O que é Visualização (de dados)
- Breve histórico
- Visualização nos dias atuais**
- Relação com outras disciplinas
- O *pipeline* de Visualização
- O papel da percepção em Visualização

3 Objetivos de uma visualização

4 Referências

Visualização nos dias atuais

Visualizações (distorcidas) de mapas de metrôs são muito utilizadas

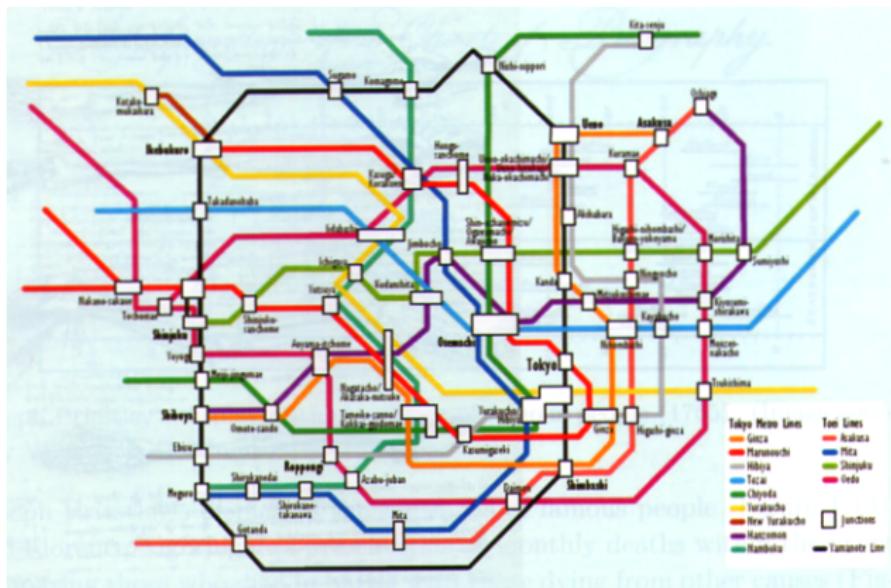


Figura: Mapa do metrô de Tóquio.

Visualização nos dias atuais

- Uma declaração como “o Dow Jones atingiu 28.000 pontos hoje” dá uma informação única e exata
- Já um gráfico dos valores médios ao longo do tempo transmite diversos elementos de informação imprecisos

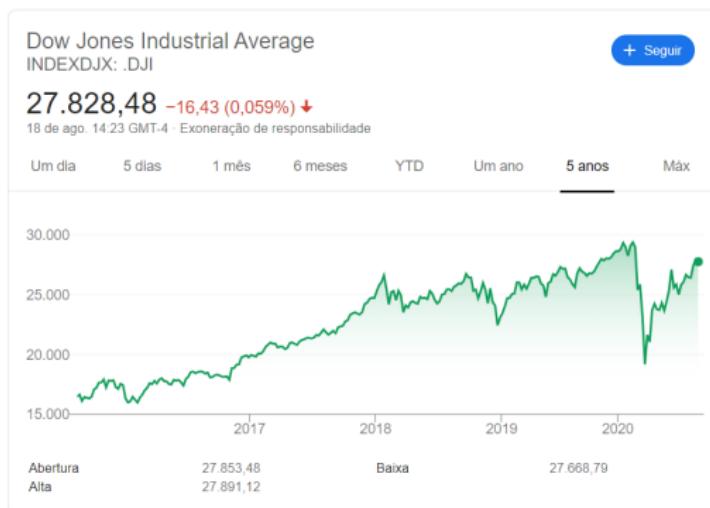


Figura: Down Jones Industrial Average ao longo de 5 anos.

Visualização nos dias atuais

Torna relativamente simples interpretar dados complexos, como eletrocardiogramas

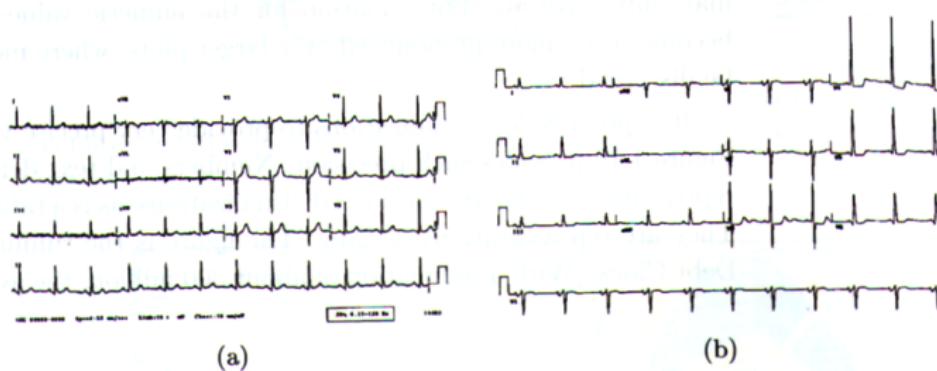


Figura: (a) eletrocardiograma de paciente adulto normal. (b) eletrocardiograma de paciente com 83 anos e problemas no coração.

Visualização nos dias atuais

Permite identificar facilmente valores espúrios, tendências e padrões, nem sempre fáceis de serem capturados com análises estatísticas

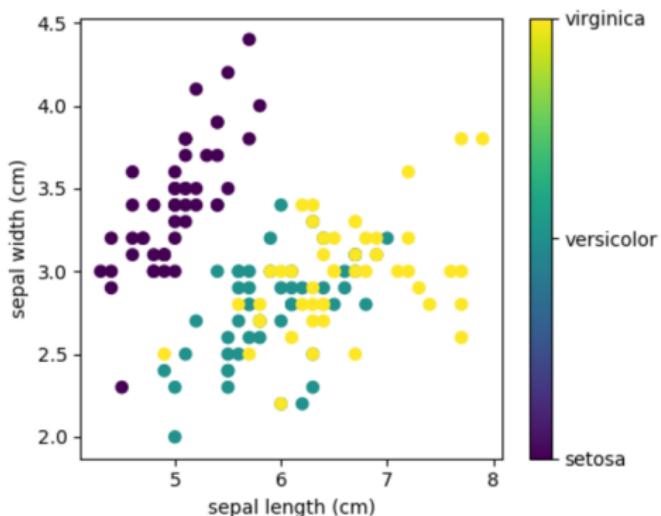


Figura: Medidas de 3 tipos de uma flor (Iris). Cor indica o tipo.

Visualização nos dias atuais

Consegue representar dados bastante complexos

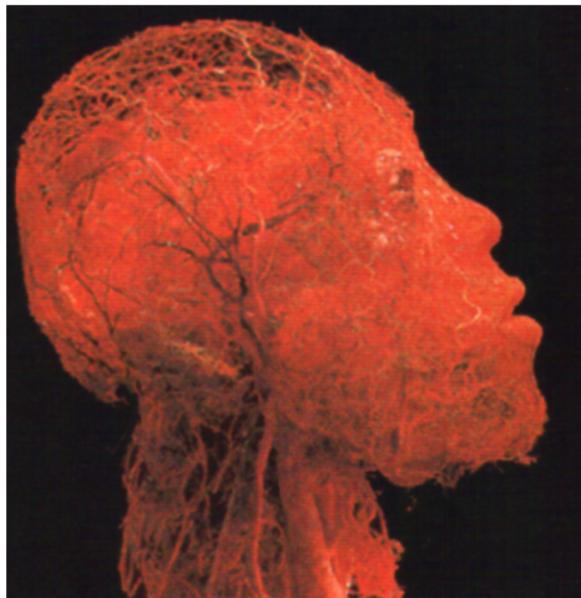


Figura: Configuração de veias na cabeça e cérebro.

Visualização nos dias atuais

- Consegue representar dados bastante complexos

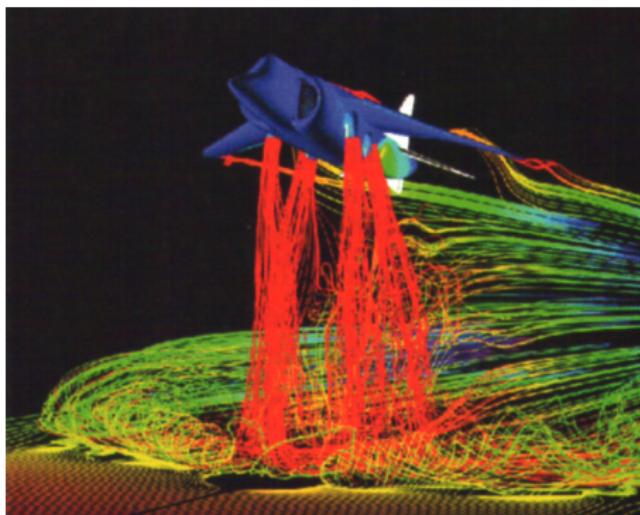


Figura: Vazão do fluxo de ar gerado por um avião a jato no momento da decolagem (simulação). Cor indica a magnitude da força exercida.

Sumário

1 Motivação

2 Panorama

- O que é Visualização (de dados)
- Breve histórico
- Visualização nos dias atuais
- **Relação com outras disciplinas**
- O *pipeline* de Visualização
- O papel da percepção em Visualização

3 Objetivos de uma visualização

4 Referências

Computação Gráfica e Visualização

- **Visualização** é o uso de gráficos para **apresentar dados**: com mapear dados em **primitivas gráficas** na tela
- **Computação gráfica** refere-se tão somente ao processo de **síntese de imagens**: como criar imagens a partir da descrição geométrica de uma cena

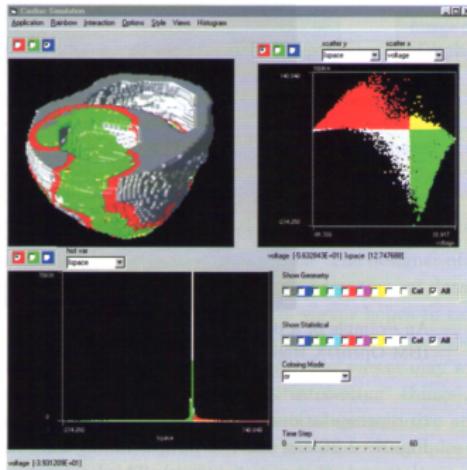


Figura: Visualização do coração de um paciente, bem como de parâmetros adicionais difíceis de representar no modelo 3D.

Visualização Científica vs. Visualização de Informação

- Diferença essencial está na natureza dos dados
 - espaciais vs. não espaciais
 - SciVis: espacialização (geometria) é parte dos dados (é dada)
 - InfoVis: espacialização é atribuída no processo de mapeamento visual (é arbitrária)

https://en.wikipedia.org/wiki/Scientific_visualization

https://en.wikipedia.org/wiki/Information_visualization

Sumário

1 Motivação

2 Panorama

- O que é Visualização (de dados)
- Breve histórico
- Visualização nos dias atuais
- Relação com outras disciplinas
- **O pipeline de Visualização**
- *O papel da percepção em Visualização*

3 Objetivos de uma visualização

4 Referências

O Pipeline de Visualização

- O pipeline de visualização define como é feito o **mapeamento dos dados em elementos gráficos**, a serem desenhados na tela
 - **Interação** tem papel fundamental nesse processo
 - Visualização integra um processo mais amplo (**de descoberta de informação/conhecimento** a partir dos dados)

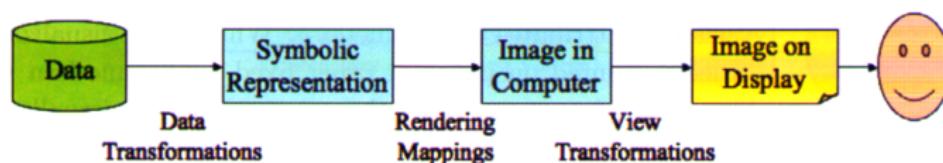


Figura: Pipeline de visualização genérico/abstrato.

Pipeline de Computação Gráfica

- O pipeline de **Computação Gráfica** visa unicamente a **síntese de imagens**
 - Modelagem
 - Viewing*
 - Recorte, remoção de superfícies ocultas
 - Projeção 3D->2D
 - Rendering*, ou exibição da cena

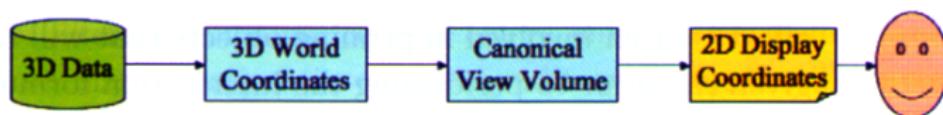


Figura: Típico pipeline de computação gráfica.

Pipeline de Visualização

- Embora similar, o **pipeline de visualização** apresenta estágios distintos
 - Organização dos dados
 - Seleção dos dados
 - Mapeamento visual dos dados
 - Definição dos parâmetros da cena
 - *Rendering*, ou geração da visualização

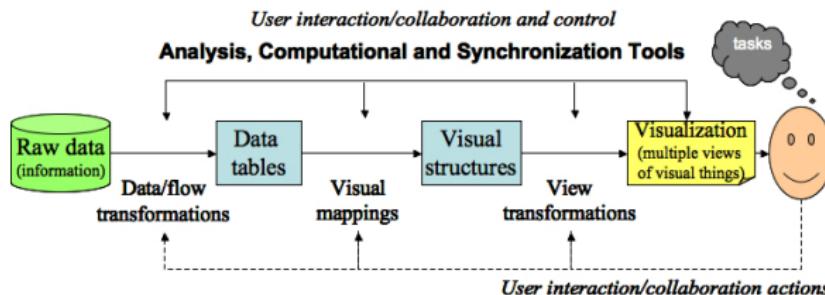
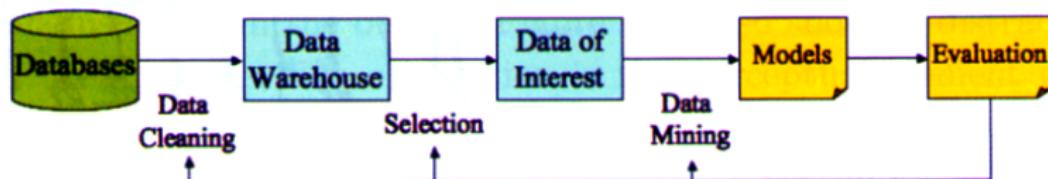


Figura: Exemplo de pipeline de visualização.

Pipeline de Mineração de Dados

- Mineração de Dados também consiste de um **pipeline**
 - Integração, limpeza, armazenamento e seleção dos dados
 - Mineração dos dados (algoritmos para identificar padrões)
 - Avaliação dos padrões
 - Avaliação dos resultados (pode visualizar!)



Sumário

1 Motivação

2 Panorama

- O que é Visualização (de dados)
- Breve histórico
- Visualização nos dias atuais
- Relação com outras disciplinas
- O *pipeline* de Visualização
- **O papel da percepção em Visualização**

3 Objetivos de uma visualização

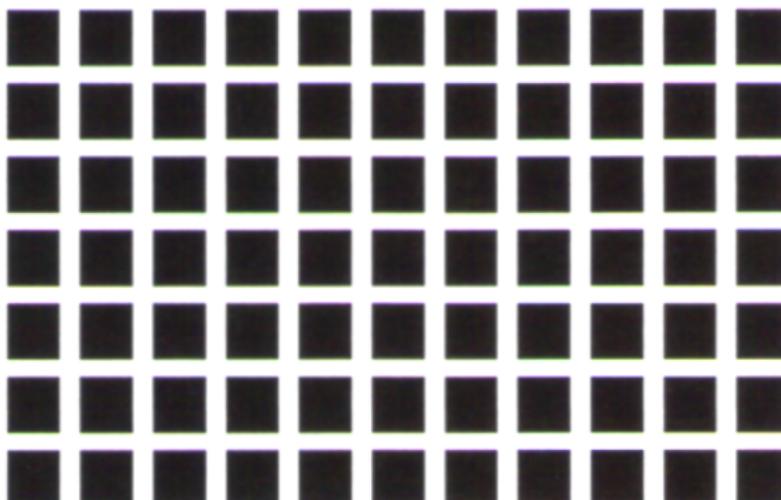
4 Referências

Percepção visual

- As **habilidades e limitações do sistema visual humano** constituem um **componente crítico** na visualização
 - Exibir gráficos atrativos é importante, mas é preciso evitar **ambiguidades** e artefatos visuais em ambientes de tomada de decisão

Percepção visual

- As **habilidades e limitações do sistema visual humano** constituem um **componente crítico** na visualização
 - Exibir gráficos atrativos é importante, mas é preciso evitar **ambiguidades** e artefatos visuais em ambientes de tomada de decisão

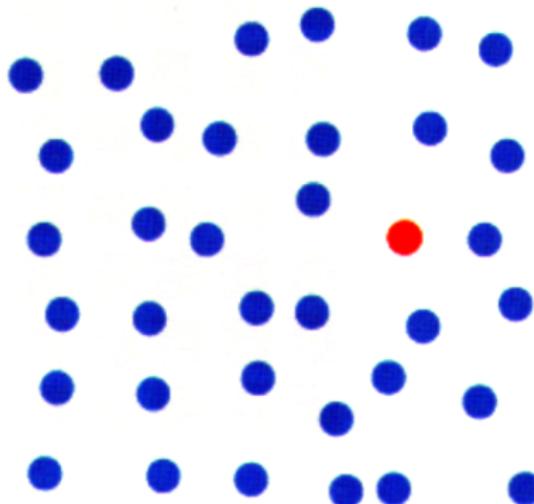


Percepção visual

- O sistema perceptual do olho humano **processa dados** de várias maneiras
 - **Processo pré-atentivo:** sistema de alto desempenho que identifica rapidamente diferenças em cor e textura, orientação de linhas, tamanho e largura de objetos, curvatura, etc.

Percepção visual

- O sistema perceptual do olho humano **processa dados** de várias maneiras
 - **Processo pré-atentivo:** sistema de alto desempenho que identifica rapidamente diferenças em cor e textura, orientação de linhas, tamanho e largura de objetos, curvatura, etc.

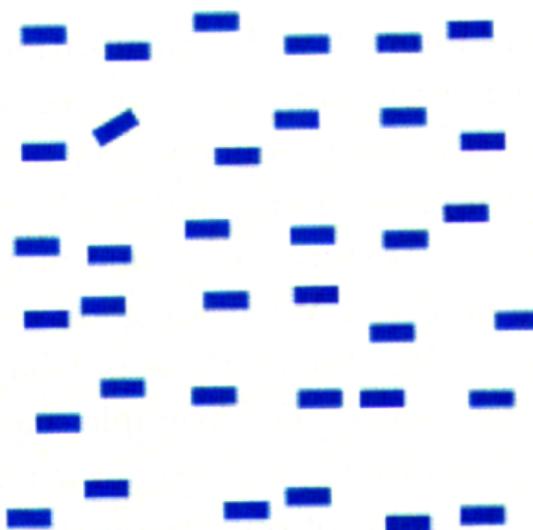


Percepção visual

- O sistema perceptual do olho humano **processa dados** de várias maneiras
 - **Processo pré-atentivo:** sistema de alto desempenho que identifica rapidamente diferenças em cor e textura, orientação de linhas, tamanho e largura de objetos, curvatura, etc.

Percepção visual

- O sistema perceptual do olho humano **processa dados** de várias maneiras
 - Processo pré-atentivo:** sistema de alto desempenho que identifica rapidamente diferenças em cor e textura, orientação de linhas, tamanho e largura de objetos, curvatura, etc.

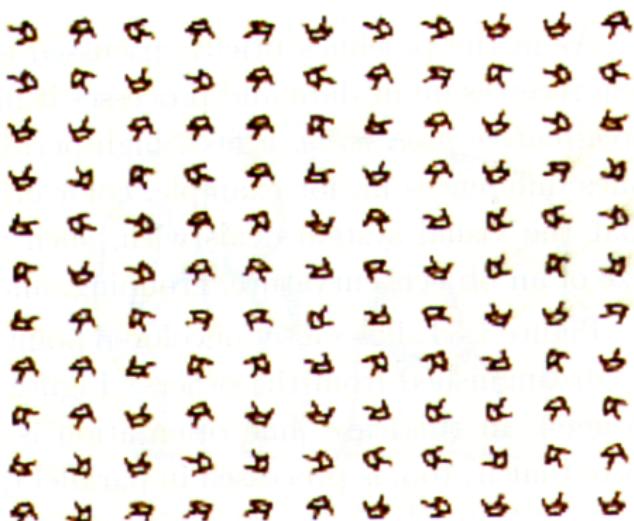


Percepção visual

Porém, outros padrões demandam esforço e atenção para serem identificados

Percepção visual

Porém, outros padrões demandam esforço e atenção para serem identificados



Percepção visual

Porém, outros padrões demandam esforço e atenção para serem identificados

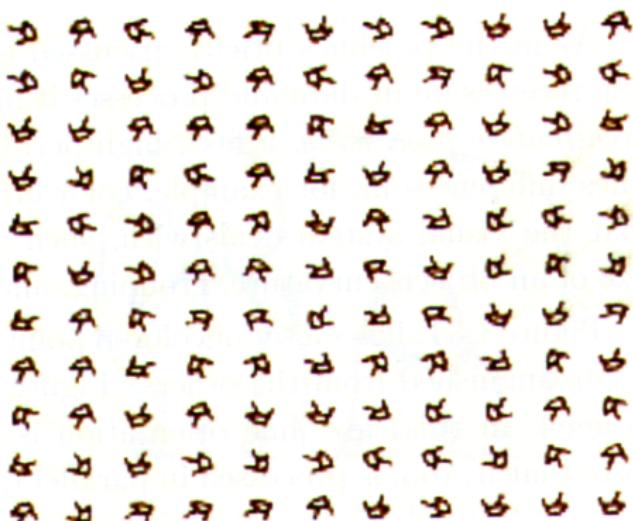


Figura: Há um quadrado de R's com orientação contrária

Percepção visual

- Uma boa compreensão do que é possível perceber visualmente com pouco esforço é de extrema importância em visualização
- Segundo a *Gestalt School of Psychology*, existem certos princípios que governam a percepção de padrões pelas pessoas:
 - Proximidade
 - Similaridade
 - Continuidade
 - Fechamento
 - Simetria
 - Plano de fundo
 - Plano de frente
 - Tamanho
- Veja
<https://www.usertesting.com/blog/gestalt-principles>

Sumário

1 Motivação

2 Panorama

- O que é Visualização (de dados)
- Breve histórico
- Visualização nos dias atuais
- Relação com outras disciplinas
- O *pipeline* de Visualização
- O papel da percepção em Visualização

3 Objetivos de uma visualização

4 Referências

Objetivos da Visualização

Pode-se criar visualizações com diferentes objetivos

- **Exploração:** pessoa quer entender, interpretar, levantar hipóteses sobre os dados
- **Confirmação:** pessoa já tem uma hipótese(s), quer verificar-a(s)
- **Apresentação:** pessoa quer apresentar um conceito, um fato ou um conjunto de conceitos/fatos para um público

O “Design space” da visualização

- Em qualquer caso, o espaço de possíveis soluções para exibir dados é imenso...
- Muitas opções, mas poucas soluções de fato adequadas
- Disciplina: conhecer um pouco sobre as opções de solução e entender quais são adequadas para que situação
- Foco: visualizações interativas, para executar tarefas de natureza exploratória ou confirmatória

O ‘Design space’ da Visualização

Limitações inerentes: computacionais, humanas, dos dispositivos de exibição

- Computacionais: tempo de processamento e memória
- Humanas: funcionamento do sistema perceptual, capacidade de atenção, de memória
- Dispositivos de exibição: número finito de pixels, densidade de informação

Sumário

1 Motivação

2 Panorama

- O que é Visualização (de dados)
- Breve histórico
- Visualização nos dias atuais
- Relação com outras disciplinas
- O *pipeline* de Visualização
- O papel da percepção em Visualização

3 Objetivos de uma visualização

4 Referências

Referências

- Ward, Matthew. Grinstein, Georges G. Keim, Daniel. **Interactive data visualization foundations, techniques, and applications.** Natick, Mass., A K Peters, 2010 (há nova edição de 2015).
 - <http://www.idvbook.com/>
- Tamara Munzner, Visualization Analysis & Design, CRC Press, 2015.