

# Introdução à Visualização

SCC5836/SCC0252 – Visualização Computacional

Profa. Maria Cristina

sala4-205

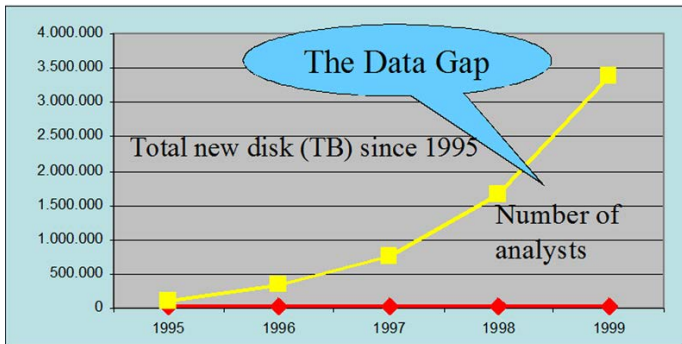
*cristina@icmc.usp.br*

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

**VICG** Grupo de Visualização,  
 **Imagens e Computação Gráfica**

- 1 **Motivação**
- 2 **Introdução**
  - O que é Visualização
  - História da Visualização
  - Visualização nos Dias Atuais
  - Relacionamento entre Visualização e outras áreas do conhecimento
  - O Processo de Visualização
- 3 **Exemplo de análise**
- 4 **Papel do Usuário**
- 5 **O Problema Fundamental**
- 6 **Sobre os slides do curso**
- 7 **Atividades**
- 8 **Referências**

# Introdução



- A informação vem crescendo 30% ao ano [Tan et. al, 2005]

# Crescimento em Números

- Em 2007 existiam cerca de 30 milhões de câmeras de vigilância somente nos EUA
  - Mais de 4 bilhões de horas de vídeo toda semana [**J. Vlahos, 2008**]

# Crescimento em Números

- Em 2007 existiam cerca de 30 milhões de câmeras de vigilância somente nos EUA
  - Mais de 4 bilhões de horas de vídeo toda semana [**J. Vlahos, 2008**]
- 2002: 5 hexabytes de nova informação em mídia impressa, magnética e ótica gerada
  - Equivalente a 37,000 cópias de todos os 7 milhões de livros da Biblioteca do Congresso americano [**Lyman & Hal, 2003**]

# Crescimento em Números

- Em 2007 existiam cerca de 30 milhões de câmeras de vigilância somente nos EUA
  - Mais de 4 bilhões de horas de vídeo toda semana [**J. Vlahos, 2008**]
- 2002: 5 hexabytes de nova informação em mídia impressa, magnética e ótica gerada
  - Equivalente a 37,000 cópias de todos os 7 milhões de livros da Biblioteca do Congresso americano [**Lyman & Hal, 2003**]
- Em uma grande empresa uma pessoa troca, em média, cerca de 177 mensagens por dia [**Tanaka, 1998**]

# Crescimento em Números

- Em 2007 existiam cerca de 30 milhões de câmeras de vigilância somente nos EUA
  - Mais de 4 bilhões de horas de vídeo toda semana [**J. Vlahos, 2008**]
- 2002: 5 hexabytes de nova informação em mídia impressa, magnética e ótica gerada
  - Equivalente a 37,000 cópias de todos os 7 milhões de livros da Biblioteca do Congresso americano [**Lyman & Hal, 2003**]
- Em uma grande empresa uma pessoa troca, em média, cerca de 177 mensagens por dia [**Tanaka, 1998**]
- Uma única edição atual do *New York Times* contém mais informação do que uma pessoa comum no século XVII teve contato em toda a sua vida [**Tanaka, 1998**]

Enfim...

Como lidar com isso?



# Sumário

- 1 Motivação
- 2 **Introdução**
  - O que é Visualização
  - História da Visualização
  - Visualização nos Dias Atuais
  - Relacionamento entre Visualização e outras áreas do conhecimento
  - O Processo de Visualização
- 3 Exemplo de análise
- 4 Papel do Usuário
- 5 O Problema Fundamental
- 6 Sobre os slides do curso
- 7 Atividades
- 8 Referências

# Sumário

- 1 Motivação
- 2 **Introdução**
  - O que é Visualização
  - História da Visualização
  - Visualização nos Dias Atuais
  - Relacionamento entre Visualização e outras áreas do conhecimento
  - O Processo de Visualização
- 3 Exemplo de análise
- 4 Papel do Usuário
- 5 O Problema Fundamental
- 6 Sobre os slides do curso
- 7 Atividades
- 8 Referências

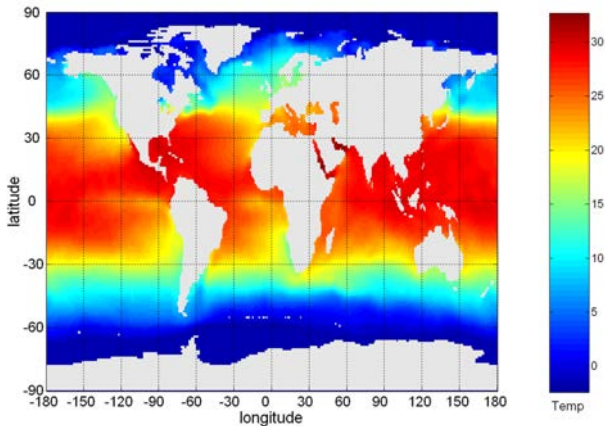
# O que é Visualização

- “Visualização é a comunicação de informação usando representações gráficas” [**Ward et al., 2010**]

# O que é Visualização

- “Visualização é a comunicação de informação usando representações gráficas” [Ward et al., 2010]
- Uma **única imagem** pode conter uma **grande quantidade de informação** e ser interpretada muito mais rapidamente do que texto
  - **Interpretação de imagens** é realizada **paralelamente** no sistema perceptual, texto é sequencial (leitura)
  - Imagem também independe da linguagem

# O que é Visualização



**Figura:** 10.000 medidas de temperatura da superfície do oceano são resumidas em uma única figura.

- **Visualização** já vem sendo empregada em diversas atividades em **substituição** à divulgação de **informação** de forma **verbal** ou **escrita**

# Visualização no Dia-a-Dia

- **Visualização** já vem sendo empregada em diversas atividades em **substituição** à divulgação de **informação** de forma **verbal ou escrita**
- Atividades regulares
  - Mapas de trem e metrô
  - Mapa de de uma região para determinar rota
  - Gráficos explicativos em jornais e revistas
  - Gráficos de previsão do tempo
  - Imagens de tomógrafos computadorizados
  - Manuais de instrução para montagem de móveis, bicicletas, etc.
- Atividades industriais
  - Análise do mercado de ações
  - Desenhos de engenharia mecânica e civil
  - Diagnóstico de câncer de mama
  - Simulação de processos complexos

# Por que Visualização é Importante?

- Gerar estatísticas dos dados não seria suficiente?
- A forma de apresentação dos dados influencia um processo de tomada de decisão?
  - Pode modificar uma decisão?
  - Pode induzir decisões erradas?
  - Existe alguma representação melhor?



# Por que Visualização é Importante?

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	

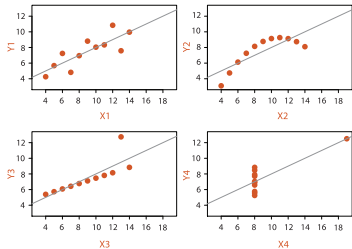
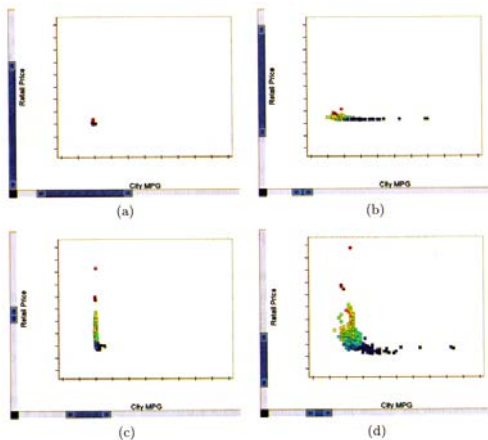


Figura: Fonte: T. Munzner Visualization Analysis & Design (Fig. 1.3).

# Por que Visualização é Importante?



**Figura:** Mesmos conjunto de dados (carros), exibido com eixos adotando diferentes escalas. (a) escala uniformemente grande em  $x$  e  $y$ . (b) escala maior em  $y$ . (c) escala maior em  $x$ . (d) escala determinada pelos intervalos de variação de  $x$  e  $y$ .

# Por que Visualização é Importante?

- A forma de apresentação dos dados tem grande impacto no processo de análise!

# Por que Visualização é Importante?

- A forma de apresentação dos dados tem grande impacto no processo de análise!

## Estudo de Caso: Comparação de Tratamentos Clínicos

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

# Por que Visualização é Importante?

Table

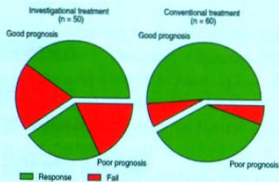
	Conventional treatment		Investigational treatment	
	Total no	% Fail	Total no	% Fail
Good prognosis	30	30	35	11
Poor prognosis	20	45	25	12
Total	50	38	60	12

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

Bar graph



Pie chart



Icon



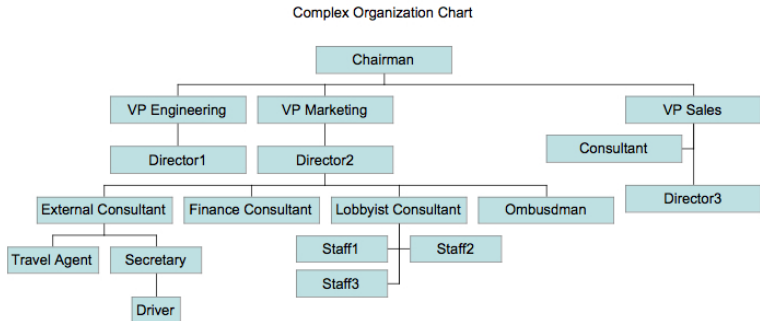
# Por que Visualização é Importante?



- 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 Visualização é Importante?

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



# Por que Visualização é Importante?

- A **importância** está em **interpretar** dados mais **rapidamente**
- Canal visual humano é muito eficiente para captar informação relevante
- Auxílio essencial para apoiar processos de descoberta de **conhecimento** e **tomada de decisão** por humanos



- 1 Motivação
- 2 **Introdução**
  - O que é Visualização
  - **História da Visualização**
  - Visualização nos Dias Atuais
  - Relacionamento entre Visualização e outras áreas do conhecimento
  - O Processo de Visualização
- 3 Exemplo de análise
- 4 Papel do Usuário
- 5 O Problema Fundamental
- 6 Sobre os slides do curso
- 7 Atividades
- 8 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

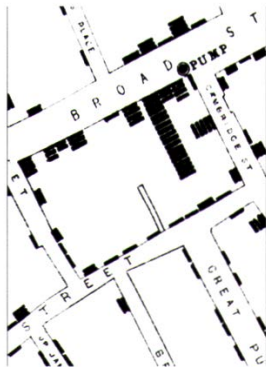
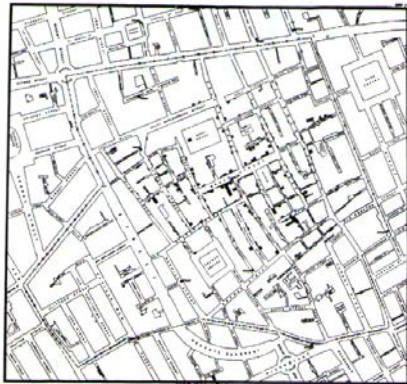
- Mapas do mundo já eram desenhados na Idade Média. Nesse, Jerusalém está no centro do mundo (catedral de Hereford, País de Gales)



Figura: Hereford Map.

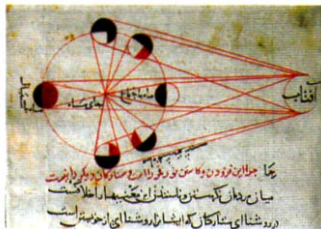
# Primeiras Visualizações

- Mapa de John Snow detalhando as mortes por cólera em Londres (em 1663)
  - Mais de 500 mortes verificadas na região de poço na *Broad Street*
  - Cada barra indica uma morte no local

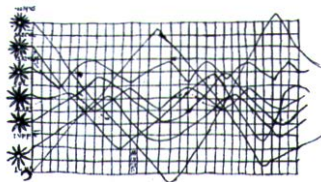


# Primeiras Visualizações

- Visualizações de séries temporais existem desde bem antes de Descartes



(a)



(b)

**Figura:** (a) Representação das fases da lua em órbita, ano de 1030. (b) movimentação dos planetas.





# Primeiras Visualizações

- Introdução do conceito de eixos representou um marco importante
- Viabiliza utilizar quaisquer variáveis como coordenadas (abstração), sem ficar restrito a interpretações geoespaciais...

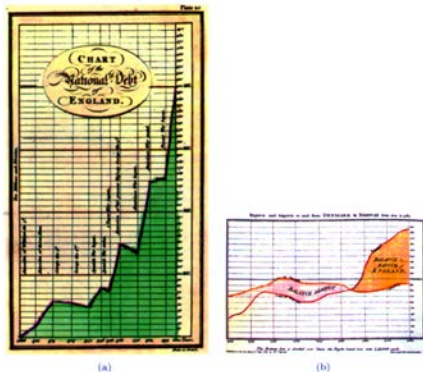


Figura: (a) Representação da evolução temporal da dívida dos EUA. (b) uma visão da balança comercial entre Inglaterra e Noruega/Dinamarca (William Playfair, 1786).

# Primeiras Visualizações

- Representação taxa de mortalidade mensal no Exército americano (causas)
  - Em azul, mortes por doenças, em vermelho, mortes por ferimentos (batalha), em preto mortes por outras causas

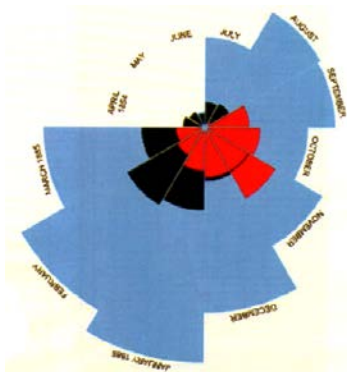


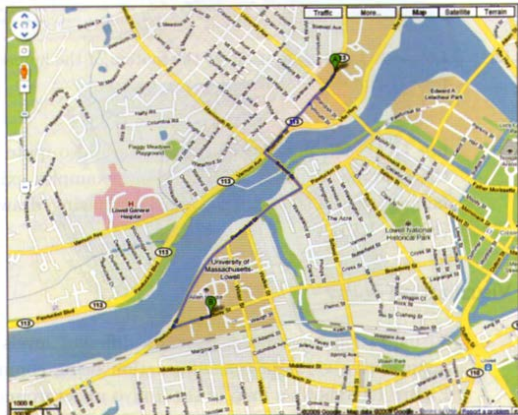
Figura: Coxcomb chart, Florence Nightingale

- 1 Motivação
- 2 **Introdução**
  - O que é Visualização
  - História da Visualização
  - **Visualização nos Dias Atuais**
  - Relacionamento entre Visualização e outras áreas do conhecimento
  - O Processo de Visualização
- 3 Exemplo de análise
- 4 Papel do Usuário
- 5 O Problema Fundamental
- 6 Sobre os slides do curso
- 7 Atividades
- 8 Referências



# Visualização Atualmente

- As distorções em um mapa (projeção 3D em 2D) são pequenas por considerarem pequenas áreas



# Visualização Atualmente

- Uma declaração como “a média da *Dow Jones* atingiu 125 pontos hoje” dá uma informação única e exata
- Já um gráfico dos valores médios transmite diversos elementos imprecisos de informação

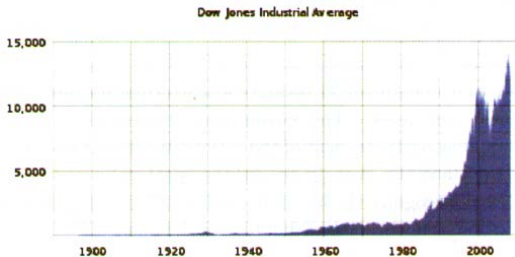
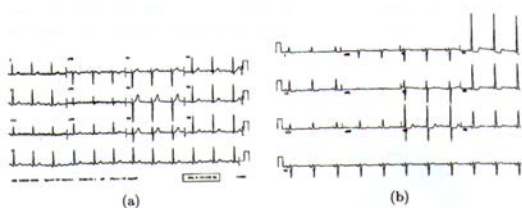


Figura: Dow Jones Industrial Average, no período 1900-2008(?).

# Visualização Atualmente

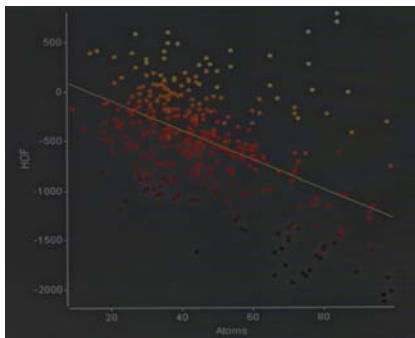
- Torna relativamente simples interpretar dados complexos, como eletrocardiogramas



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

# Visualização Atualmente

- Permite identificar facilmente valores espúrios, tendências e padrões difíceis de serem capturados com análises estatísticas



**Figura:** Dados sobre mecanismo da ação de leveduras (fermento), com linha de regressão. HOF indica calor de formação.



# Visualização Atualmente

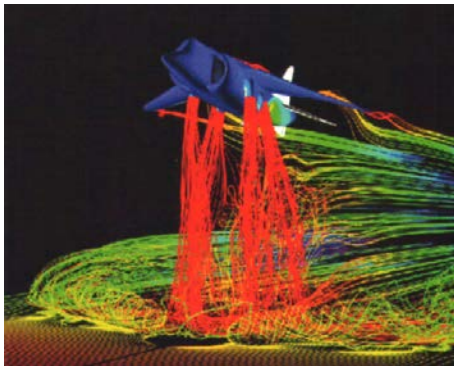
- Consegue representar dados bastante complexos



Figura: Configuração de veias na cabeça e cérebro.

# Visualização Atualmente

- Consegue representar dados bastante complexos



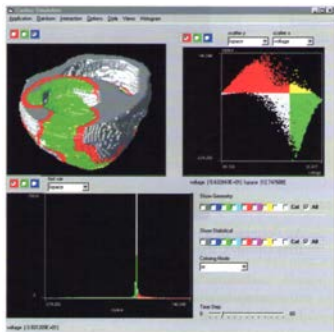
**Figura:** Simulação da vazão do ar gerado por um avião a jato no momento da decolagem. Cor indica a quantidade de força exercida.

# Sumário

- 1 Motivação
- 2 **Introdução**
  - O que é Visualização
  - História da Visualização
  - Visualização nos Dias Atuais
  - **Relacionamento entre Visualização e outras áreas do conhecimento**
  - O Processo de Visualização
- 3 Exemplo de análise
- 4 Papel do Usuário
- 5 O Problema Fundamental
- 6 Sobre os slides do curso
- 7 Atividades
- 8 Referências

# Diferença entre Computação Gráfica e Visualização

- **Visualização** é a aplicação de gráficos para **apresentar dados**, que são mapeados na tela como **primitivas gráficas**
  - Computação gráfica abrange não somente os processos 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 serem representados no modelo 3D.

# Visualização Científica vs. Visualização de Informação

- Uma diferença essencial está na natureza dos dados
  - espaciais vs não espaciais
  - SciVis: posição espacial é parte dos dados (é dada)
  - InfoVis: posição espacial é atribuída no processo de mapeamento visual (é arbitrária)

- 1 Motivação
- 2 **Introdução**
  - O que é Visualização
  - História da Visualização
  - Visualização nos Dias Atuais
  - Relacionamento entre Visualização e outras áreas do conhecimento
  - **O Processo de Visualização**
- 3 Exemplo de análise
- 4 Papel do Usuário
- 5 O Problema Fundamental
- 6 Sobre os slides do curso
- 7 Atividades
- 8 Referências

# O Processo de Visualização

- O processo de visualização é caracterizado por definir um **mapeamento dos dados para elementos gráficos**, que são então desenhados na tela
  - **Interação** tem papel fundamental nesse processo
  - Visualização integra um processo mais amplo (**descoberta de conhecimento**)

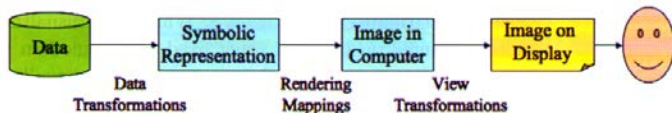


Figura: Processo de visualização genérico/abstrato.

# 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

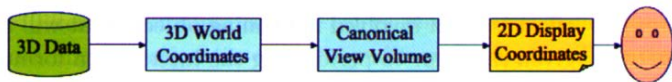


Figura: Típico pipeline de computação gráfica.



# Pipeline de Visualização

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

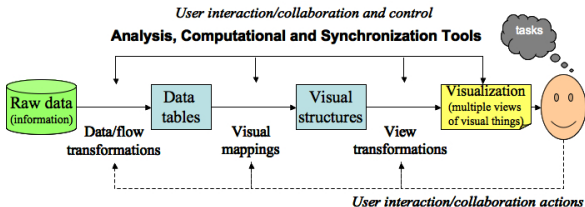
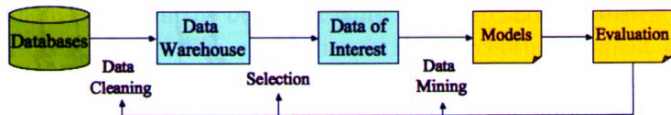


Figura: Um exemplo de pipeline de visualização.

# Pipeline de Descoberta de Conhecimento

- **Descoberta de conhecimento** (Mineração de Dados) também ocorre como um **pipeline** e inclui os seguintes passos
  - Dados
  - Integração, limpeza, armazenamento e seleção dos dados
  - Mineração dos dados
  - Avaliação dos padrões
  - Visualização (dos resultados)



# O Papel da Percepção

- Na visualização, um **componente crítico** são as **habilidades e limitações do sistema visual humano**
  - Exibir gráficos atrativos é importante, mas **ambiguidades** (ou ilusões) precisam ser **evitadas** em ambientes de tomada de decisão

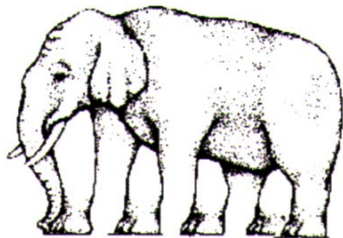
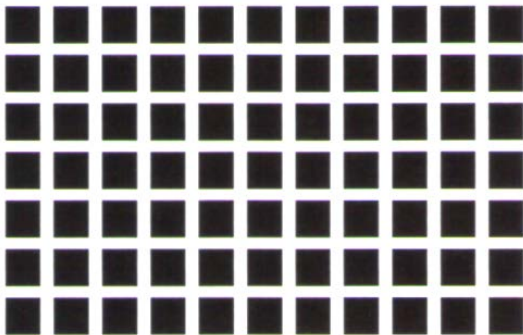


Figura: Quantas pernas tem o elefante?

- <http://www.ritsumei.ac.jp/~akitaoka/index-e.html>

# O Papel da Percepção

- **Artefatos visuais:** cuidado para **não mapear** uma variável em atributos gráficos que temos **habilidade limitada** para controlar ou quantificar
  - Ex. de atributos gráficos: textura, cor, comprimento, área, movimento
  - Texturas são ruins, p.ex., para representar valores numéricos



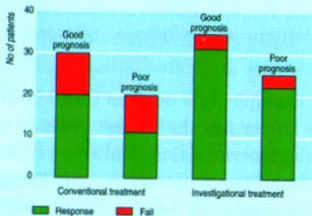
# Por que Visualização é Importante?

Table

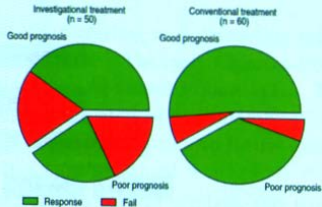
	Conventional treatment		Investigational treatment	
	Total no	% Fail	Total no	% Fail
Good prognosis	30	30	35	11
Poor prognosis	20	45	25	12
Total	50	38	60	12

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

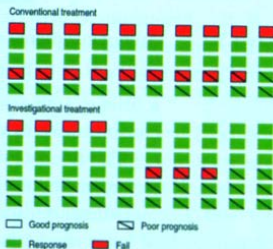
Bar graph



Pie chart



Icon

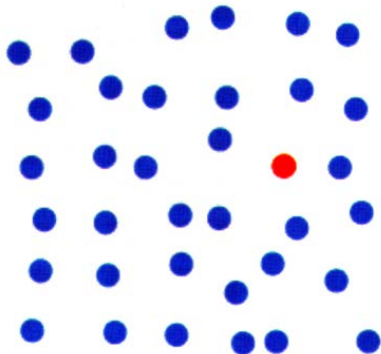


# O Papel da Percepção

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

# O Papel da Percepção

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



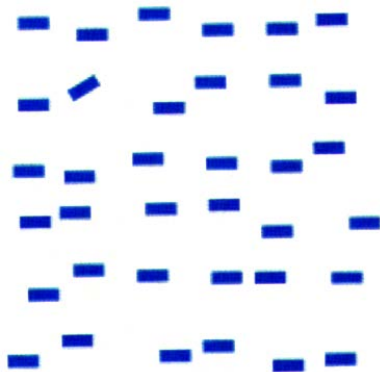
# O Papel da Percepção

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



# O Papel da Percepção

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

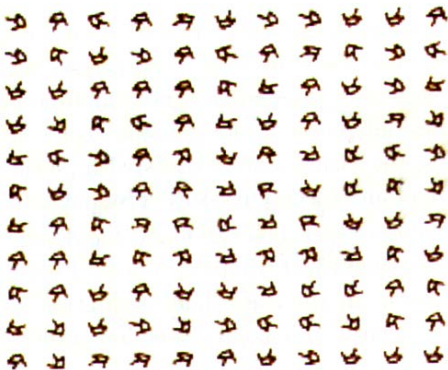


# O Papel da Percepção

- Porém, alguns padrões demandam atenção para serem identificados

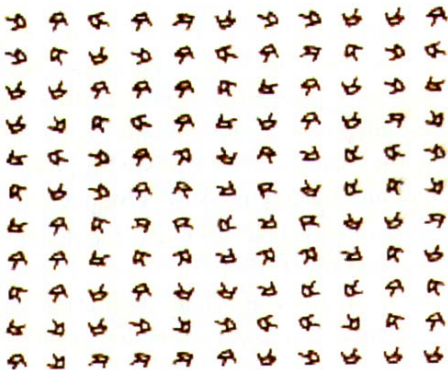
# O Papel da Percepção

- Porém, alguns padrões demandam atenção para serem identificados



# O Papel da Percepção

- Porém, alguns padrões demandam atenção para serem identificados



- Há um quadrado de R's com orientação contrária

# O Papel da Percepção

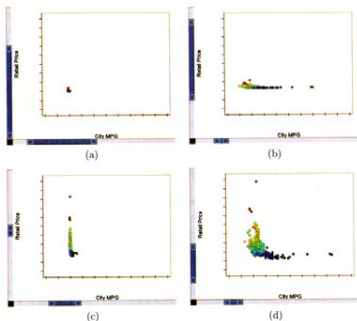
- Uma boa compreensão do que é possível perceber visualmente é de extrema importância em visualização
- Segundo a *Gestalt School of Psychology*, as leis pelas quais percebemos padrões são
  - Proximidade
  - Similaridade
  - Continuidade
  - Fechamento
  - Simetria
  - Plano de fundo
  - Plano de frente
  - Tamanho
- ver <http://graphicdesign.spokanefalls.edu/tutorials/process/gestaltprinciples/gestaltprinc.htm>

# Sumário

- 1 Motivação
- 2 Introdução
  - O que é Visualização
  - História da Visualização
  - Visualização nos Dias Atuais
  - Relacionamento entre Visualização e outras áreas do conhecimento
  - O Processo de Visualização
- 3 Exemplo de análise**
- 4 Papel do Usuário
- 5 O Problema Fundamental
- 6 Sobre os slides do curso
- 7 Atividades
- 8 Referências

# Scatterplots

- Gráficos de Dispersão
- Uma das representações visuais mais antigas
  - permite comparar dois atributos (variáveis): cada instância de dado mapeado em um elemento gráfico, cuja posição espacial é determinada pelos valores dos respectivos atributos
  - Dois eixos perpendiculares entre si
  - Outros atributos podem ser mapeados para a cor e/ou o tamanho dos marcadores gráficos



# Estudo de Caso

## Estudo de Caso

- Conjunto de dados de carros e caminhões (428 veículos)

Vehicle Name	Small/Sporty/ Compact/Large		Sports Car	SUV	Wagon	Minivan	Pickup	AWD	RWD	Retail Price	Dealer Cost	Engine Size (l)	Cyl	HP	City MPG	Hwy MPG	Weight	Wheel		
	Sedan																	Base	Len	Width
Toyota 4Runner SR5 V6			0	0	1	0	0	0	0	27710	24801	4	6	245	18	21	4035	110	189	74
Toyota Avalon XL 4dr	1	0	0	0	0	0	0	0	0	26560	23693	3	6	210	21	29	3417	107	192	72
Toyota Avalon XLS 4dr	1	0	0	0	0	0	0	0	0	30920	27271	3	6	210	21	29	3439	107	192	72
Toyota Camry LE 4dr	1	0	0	0	0	0	0	0	0	19560	17558	2.4	4	157	24	33	3086	107	189	71
Toyota Camry LE V6 4dr	1	0	0	0	0	0	0	0	0	22775	20325	3	6	210	21	29	3296	107	189	71
Toyota Camry Solara SE 2dr	1	0	0	0	0	0	0	0	0	19635	17722	2.4	4	157	24	33	3175	107	193	72
Toyota Camry Solara SE V6 2dr	1	0	0	0	0	0	0	0	0	21965	19819	3.3	6	225	20	29	3417	107	193	72
Toyota Camry Solara SLE V6 2dr	1	0	0	0	0	0	0	0	0	26510	23908	3.3	6	225	20	29	3439	107	193	72
Toyota Camry XLE V6 4dr	1	0	0	0	0	0	0	0	0	25920	23125	3	6	210	21	29	3362	107	189	71
Toyota Celica GT-S 2dr	0	1	0	0	0	0	0	0	0	22570	20363	1.8	4	180	24	33	2500	102	171	68
Toyota Corolla CE 4dr	1	0	0	0	0	0	0	0	0	14085	13065	1.8	4	130	32	40	2502	102	178	67
Toyota Corolla LE 4dr	1	0	0	0	0	0	0	0	0	15295	13889	1.8	4	130	32	40	2524	102	178	67
Toyota Corolla S 4dr	1	0	0	0	0	0	0	0	0	15030	13650	1.8	4	130	32	40	2524	102	178	67
Toyota Echo 2dr auto	1	0	0	0	0	0	0	0	0	11560	10896	1.5	4	108	33	39	2085	93	163	65
Toyota Echo 2dr manual	1	0	0	0	0	0	0	0	0	10760	10144	1.5	4	108	35	43	2035	93	163	65
Toyota Echo 4dr	1	0	0	0	0	0	0	0	0	11290	10642	1.5	4	108	35	43	2055	93	163	65
Toyota Highlander V6	0	0	1	0	0	0	0	1	0	27930	24915	3.3	6	230	18	24	3935	107	185	72
Toyota Land Cruiser	0	0	1	0	0	0	0	1	0	54765	47986	4.7	8	325	13	17	5390	112	193	76
Toyota Matrix XR	0	0	0	1	0	0	0	0	0	16695	15156	1.8	4	130	29	36	2679	102	171	70
Toyota MR2 Spyder convertible 2dr	0	1	0	0	0	0	0	0	0	25130	22787	1.8	4	138	26	32	2195	97	153	67
Toyota Prius 4dr (gas/electric)	1	0	0	0	0	0	0	0	0	20510	18926	1.5	4	110	59	51	2890	106	175	68
Toyota RAV4	0	0	1	0	0	0	0	1	0	20290	18553	2.4	4	161	22	27	3119	98	167	68
Toyota Sequoia SR5	0	0	1	0	0	0	0	1	0	35695	31827	4.7	8	240	14	17	5270	118	204	76
Toyota Sienna CE	0	0	0	0	1	0	0	0	0	23495	21198	3.3	6	230	19	27	4120	119	200	77
Toyota Sienna XLE Limited	0	0	0	0	1	0	0	0	0	28800	25690	3.3	6	230	19	27	4165	119	200	77
Toyota Tacoma	0	0	0	0	0	0	1	0	1	12800	11879	2.4	4	142	22	27	2750	103	*	*
Toyota Tundra Access Cab V6 SR5	0	0	0	0	0	0	1	1	0	25935	23520	3.4	6	190	14	17	4435	128	*	*
Toyota Tundra Regular Cab V6	0	0	0	0	0	0	1	0	1	16495	14978	3.4	6	190	16	20	3925	128	*	*

Figura: Somente os veículos Toyota (28 instâncias).



# Estudo de Caso

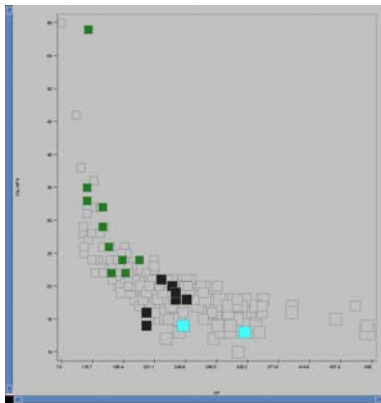
## Perguntas

- Qual a relação entre o modelo do veículo e o consumo de combustível?
- Há relação entre preço de venda do carro e seu consumo?
- Veículos de outras nacionalidades são mais econômicos que os americanos?

Vehicle Name	Manufacturer			Sales				Retail			Dealer			Engine			City			High			Hwy		
	Country	Origin	Year	MSRP	MSRP	MSRP	MSRP	MSRP	MSRP	MSRP	MSRP	MSRP	MSRP	MSRP	MSRP	MSRP	MSRP	MSRP	MSRP	MSRP	MSRP	MSRP	MSRP	MSRP	
Toyota Avalon 3.5LE V6	0	0	1	0	0	0	0	0	27110	24000	4	6	249	18	21	4236	110	189	74						
Toyota Avalon XL 4dr	1	0	0	0	0	0	0	0	30660	23000	3	6	210	21	26	3417	107	182	72						
Toyota Avalon XL 4dr	1	0	0	0	0	0	0	0	30920	27271	3	6	210	21	26	3438	107	182	72						
Toyota Camry LE 4dr	1	0	0	0	0	0	0	0	19980	17088	2	4	167	24	33	3086	107	189	71						
Toyota Camry LE V6 4dr	1	0	0	0	0	0	0	0	22775	20326	3	6	210	21	26	3266	107	189	71						
Toyota Camry Solara SE 2dr	1	0	0	0	0	0	0	0	19636	17222	2	4	167	24	33	3175	107	183	72						
Toyota Camry Solara SE V6 2dr	1	0	0	0	0	0	0	0	21888	19819	3	6	225	25	29	3417	107	183	72						
Toyota Camry Solara SE V6 2dr	1	0	0	0	0	0	0	0	26110	23008	3	6	225	25	29	3438	107	183	72						
Toyota Camry SE V6 4dr	1	0	0	0	0	0	0	0	26020	23126	3	6	210	21	26	3382	107	189	71						
Toyota Corolla CE 4dr	1	0	0	0	0	0	0	0	12570	10863	1	4	180	24	33	2300	102	171	68						
Toyota Corolla CE 4dr	1	0	0	0	0	0	0	0	14085	13066	1	4	130	32	40	2362	102	176	67						
Toyota Corolla LE 4dr	1	0	0	0	0	0	0	0	16296	13999	1	4	130	32	40	2524	102	176	67						
Toyota Corolla S 4dr	1	0	0	0	0	0	0	0	16320	13662	1	4	130	32	40	2524	102	176	67						
Toyota Echo 2dr auto	1	0	0	0	0	0	0	0	11580	10996	1	4	136	33	36	2385	93	163	66						
Toyota Echo 2dr manual	1	0	0	0	0	0	0	0	10760	10144	1	4	136	35	43	2236	93	163	66						
Toyota Echo 4dr	1	0	0	0	0	0	0	0	11290	10642	1	4	168	36	43	2356	93	163	66						
Toyota Highlander V6	0	0	1	0	0	0	1	0	27930	24915	3	6	230	18	24	3836	107	186	72						
Toyota Land Cruiser	0	0	1	0	0	0	1	0	54780	47986	4	8	325	13	17	5290	112	193	76						
Toyota Matrix SE	0	0	0	1	0	0	0	0	16895	15156	1	4	130	29	36	2679	102	171	70						
Toyota MR2 Spyder convertible 2dr	0	0	1	0	0	0	0	0	12610	23797	1	4	138	26	33	2196	87	163	67						
Toyota Prius 4dr (gas/electric)	1	0	0	0	0	0	0	0	25510	19928	1	4	110	18	19	2890	108	175	68						
Toyota Matrix	0	0	1	0	0	0	1	0	32390	19953	2	4	181	22	37	3119	96	167	68						
Toyota Sienna 2005	0	0	1	0	0	0	1	0	26696	21627	4	8	240	14	17	6270	118	204	78						
Toyota Sienna 4dr (gas/electric)	1	0	0	0	0	0	0	0	25496	21186	3	6	230	19	27	4120	118	205	77						
Toyota Sienna XLE Limited	0	0	0	1	0	0	0	0	29800	26980	3	6	230	19	27	4165	118	205	77						
Toyota Tacoma	0	0	0	0	0	1	0	1	12800	11879	2	4	142	22	27	2760	93	163	66						
Toyota Tundra Access Cab V6 SR5	0	0	0	0	1	0	1	0	28935	23632	3	6	192	14	17	4426	128	211	81						
Toyota Tundra Regular Cab V6	0	0	0	0	0	1	0	1	16495	14879	3	6	190	16	20	3826	128	211	81						

Figura: Somente os veículos Toyota (28 instâncias).

# Estudo de Caso



**Figura:** Comparando a potência vs. consumo dos carros Toyota. A cor mapeia a classe do veículo (esporte, minivan, pickup).

# Estudo de Caso

- Selecionando veículos de outro fabricante, o padrão de relacionamento linear entre potência e consumo se mantém?

# Estudo de Caso

- Selecione veículos de outro fabricante, o padrão de relacionamento linear entre potência e consumo se mantém?

Veículo	Modelo	Quilômetros	Quilômetros por Litro	Velocidade	Potência	Consumo	Preço	Preço por Quilômetro
Ford Focus Sedan 1.8	Ford	15000	12.5	150	125	10	15000	15000
Ford Focus Sedan 2.0	Ford	15000	12.5	150	150	10	18000	18000
Ford Focus Sedan 2.5	Ford	15000	12.5	150	200	10	22000	22000
Ford Focus Sedan 3.0	Ford	15000	12.5	150	250	10	25000	25000
Ford Focus Sedan 4.0	Ford	15000	12.5	150	350	10	35000	35000
Ford Focus Sedan 5.0	Ford	15000	12.5	150	450	10	45000	45000
Ford Focus Sedan 6.0	Ford	15000	12.5	150	600	10	60000	60000
Ford Focus Sedan 7.0	Ford	15000	12.5	150	750	10	75000	75000
Ford Focus Sedan 8.0	Ford	15000	12.5	150	900	10	90000	90000
Ford Focus Sedan 9.0	Ford	15000	12.5	150	1050	10	105000	105000
Ford Focus Sedan 10.0	Ford	15000	12.5	150	1200	10	120000	120000
Ford Focus Sedan 11.0	Ford	15000	12.5	150	1350	10	135000	135000
Ford Focus Sedan 12.0	Ford	15000	12.5	150	1500	10	150000	150000
Ford Focus Sedan 13.0	Ford	15000	12.5	150	1650	10	165000	165000
Ford Focus Sedan 14.0	Ford	15000	12.5	150	1800	10	180000	180000
Ford Focus Sedan 15.0	Ford	15000	12.5	150	1950	10	195000	195000
Ford Focus Sedan 16.0	Ford	15000	12.5	150	2100	10	210000	210000
Ford Focus Sedan 17.0	Ford	15000	12.5	150	2250	10	225000	225000
Ford Focus Sedan 18.0	Ford	15000	12.5	150	2400	10	240000	240000
Ford Focus Sedan 19.0	Ford	15000	12.5	150	2550	10	255000	255000
Ford Focus Sedan 20.0	Ford	15000	12.5	150	2700	10	270000	270000
Ford Focus Sedan 21.0	Ford	15000	12.5	150	2850	10	285000	285000
Ford Focus Sedan 22.0	Ford	15000	12.5	150	3000	10	300000	300000
Ford Focus Sedan 23.0	Ford	15000	12.5	150	3150	10	315000	315000
Ford Focus Sedan 24.0	Ford	15000	12.5	150	3300	10	330000	330000
Ford Focus Sedan 25.0	Ford	15000	12.5	150	3450	10	345000	345000
Ford Focus Sedan 26.0	Ford	15000	12.5	150	3600	10	360000	360000
Ford Focus Sedan 27.0	Ford	15000	12.5	150	3750	10	375000	375000
Ford Focus Sedan 28.0	Ford	15000	12.5	150	3900	10	390000	390000
Ford Focus Sedan 29.0	Ford	15000	12.5	150	4050	10	405000	405000
Ford Focus Sedan 30.0	Ford	15000	12.5	150	4200	10	420000	420000
Ford Focus Sedan 31.0	Ford	15000	12.5	150	4350	10	435000	435000
Ford Focus Sedan 32.0	Ford	15000	12.5	150	4500	10	450000	450000
Ford Focus Sedan 33.0	Ford	15000	12.5	150	4650	10	465000	465000
Ford Focus Sedan 34.0	Ford	15000	12.5	150	4800	10	480000	480000
Ford Focus Sedan 35.0	Ford	15000	12.5	150	4950	10	495000	495000
Ford Focus Sedan 36.0	Ford	15000	12.5	150	5100	10	510000	510000
Ford Focus Sedan 37.0	Ford	15000	12.5	150	5250	10	525000	525000
Ford Focus Sedan 38.0	Ford	15000	12.5	150	5400	10	540000	540000
Ford Focus Sedan 39.0	Ford	15000	12.5	150	5550	10	555000	555000
Ford Focus Sedan 40.0	Ford	15000	12.5	150	5700	10	570000	570000
Ford Focus Sedan 41.0	Ford	15000	12.5	150	5850	10	585000	585000
Ford Focus Sedan 42.0	Ford	15000	12.5	150	6000	10	600000	600000
Ford Focus Sedan 43.0	Ford	15000	12.5	150	6150	10	615000	615000
Ford Focus Sedan 44.0	Ford	15000	12.5	150	6300	10	630000	630000
Ford Focus Sedan 45.0	Ford	15000	12.5	150	6450	10	645000	645000
Ford Focus Sedan 46.0	Ford	15000	12.5	150	6600	10	660000	660000
Ford Focus Sedan 47.0	Ford	15000	12.5	150	6750	10	675000	675000
Ford Focus Sedan 48.0	Ford	15000	12.5	150	6900	10	690000	690000
Ford Focus Sedan 49.0	Ford	15000	12.5	150	7050	10	705000	705000
Ford Focus Sedan 50.0	Ford	15000	12.5	150	7200	10	720000	720000
Ford Focus Sedan 51.0	Ford	15000	12.5	150	7350	10	735000	735000
Ford Focus Sedan 52.0	Ford	15000	12.5	150	7500	10	750000	750000
Ford Focus Sedan 53.0	Ford	15000	12.5	150	7650	10	765000	765000
Ford Focus Sedan 54.0	Ford	15000	12.5	150	7800	10	780000	780000
Ford Focus Sedan 55.0	Ford	15000	12.5	150	7950	10	795000	795000
Ford Focus Sedan 56.0	Ford	15000	12.5	150	8100	10	810000	810000
Ford Focus Sedan 57.0	Ford	15000	12.5	150	8250	10	825000	825000
Ford Focus Sedan 58.0	Ford	15000	12.5	150	8400	10	840000	840000
Ford Focus Sedan 59.0	Ford	15000	12.5	150	8550	10	855000	855000
Ford Focus Sedan 60.0	Ford	15000	12.5	150	8700	10	870000	870000
Ford Focus Sedan 61.0	Ford	15000	12.5	150	8850	10	885000	885000
Ford Focus Sedan 62.0	Ford	15000	12.5	150	9000	10	900000	900000
Ford Focus Sedan 63.0	Ford	15000	12.5	150	9150	10	915000	915000
Ford Focus Sedan 64.0	Ford	15000	12.5	150	9300	10	930000	930000
Ford Focus Sedan 65.0	Ford	15000	12.5	150	9450	10	945000	945000
Ford Focus Sedan 66.0	Ford	15000	12.5	150	9600	10	960000	960000
Ford Focus Sedan 67.0	Ford	15000	12.5	150	9750	10	975000	975000
Ford Focus Sedan 68.0	Ford	15000	12.5	150	9900	10	990000	990000
Ford Focus Sedan 69.0	Ford	15000	12.5	150	10050	10	1005000	1005000
Ford Focus Sedan 70.0	Ford	15000	12.5	150	10200	10	1020000	1020000
Ford Focus Sedan 71.0	Ford	15000	12.5	150	10350	10	1035000	1035000
Ford Focus Sedan 72.0	Ford	15000	12.5	150	10500	10	1050000	1050000
Ford Focus Sedan 73.0	Ford	15000	12.5	150	10650	10	1065000	1065000
Ford Focus Sedan 74.0	Ford	15000	12.5	150	10800	10	1080000	1080000
Ford Focus Sedan 75.0	Ford	15000	12.5	150	10950	10	1095000	1095000
Ford Focus Sedan 76.0	Ford	15000	12.5	150	11100	10	1110000	1110000
Ford Focus Sedan 77.0	Ford	15000	12.5	150	11250	10	1125000	1125000
Ford Focus Sedan 78.0	Ford	15000	12.5	150	11400	10	1140000	1140000
Ford Focus Sedan 79.0	Ford	15000	12.5	150	11550	10	1155000	1155000
Ford Focus Sedan 80.0	Ford	15000	12.5	150	11700	10	1170000	1170000
Ford Focus Sedan 81.0	Ford	15000	12.5	150	11850	10	1185000	1185000
Ford Focus Sedan 82.0	Ford	15000	12.5	150	12000	10	1200000	1200000
Ford Focus Sedan 83.0	Ford	15000	12.5	150	12150	10	1215000	1215000
Ford Focus Sedan 84.0	Ford	15000	12.5	150	12300	10	1230000	1230000
Ford Focus Sedan 85.0	Ford	15000	12.5	150	12450	10	1245000	1245000
Ford Focus Sedan 86.0	Ford	15000	12.5	150	12600	10	1260000	1260000
Ford Focus Sedan 87.0	Ford	15000	12.5	150	12750	10	1275000	1275000
Ford Focus Sedan 88.0	Ford	15000	12.5	150	12900	10	1290000	1290000
Ford Focus Sedan 89.0	Ford	15000	12.5	150	13050	10	1305000	1305000
Ford Focus Sedan 90.0	Ford	15000	12.5	150	13200	10	1320000	1320000
Ford Focus Sedan 91.0	Ford	15000	12.5	150	13350	10	1335000	1335000
Ford Focus Sedan 92.0	Ford	15000	12.5	150	13500	10	1350000	1350000
Ford Focus Sedan 93.0	Ford	15000	12.5	150	13650	10	1365000	1365000
Ford Focus Sedan 94.0	Ford	15000	12.5	150	13800	10	1380000	1380000
Ford Focus Sedan 95.0	Ford	15000	12.5	150	13950	10	1395000	1395000
Ford Focus Sedan 96.0	Ford	15000	12.5	150	14100	10	1410000	1410000
Ford Focus Sedan 97.0	Ford	15000	12.5	150	14250	10	1425000	1425000
Ford Focus Sedan 98.0	Ford	15000	12.5	150	14400	10	1440000	1440000
Ford Focus Sedan 99.0	Ford	15000	12.5	150	14550	10	1455000	1455000
Ford Focus Sedan 100.0	Ford	15000	12.5	150	14700	10	1470000	1470000

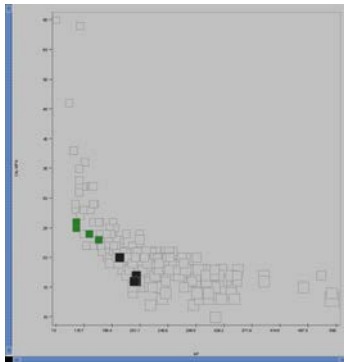
Figura: Selecionando veículos Kia...

# Estudo de Caso

- Selecionando veículos de outro fabricante, o padrão de relacionamento linear entre potência e consumo se mantém?

# Estudo de Caso

- Selecionando veículos de outro fabricante, o padrão de relacionamento linear entre potência e consumo se mantém?



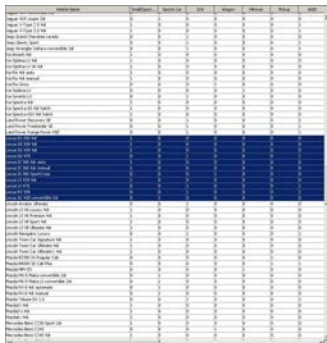
**Figura:** Scatterplot potência versus consumo para veículos Kia. A relação (aproximada) linear é mantida.

# Estudo de Caso

- Formamos uma hipótese: potência é inversamente proporcional ao consumo
  - Observada em carros de fabricantes não americanos
  - Pode ser confirmada de modo geral?

# Estudo de Caso

- Formamos uma hipótese: potência é inversamente proporcional ao consumo
  - Observada em carros de fabricantes não americanos
  - Pode ser confirmada de modo geral?



The image shows a spreadsheet with columns for 'Fabricante', 'Modelo', 'Ano', 'Potência', 'Consumo', 'Velocidade', and 'Preço'. The rows list various car models from different manufacturers. A row for a Lexus vehicle is highlighted in blue. The highlighted row is: 'Lexus LX 570 4x4'. Other visible rows include 'Lexus ES 350 4x4', 'Lexus ES 350', 'Lexus ES 300 4x4', 'Lexus ES 300', 'Lexus ES 250 4x4', 'Lexus ES 250', 'Lexus ES 240', 'Lexus ES 230', 'Lexus ES 230 4x4', 'Lexus ES 200', 'Lexus ES 180', 'Lexus ES 180 4x4', 'Lexus ES 180', 'Lexus ES 170', 'Lexus ES 170 4x4', 'Lexus ES 170', 'Lexus ES 160', 'Lexus ES 160 4x4', 'Lexus ES 160', 'Lexus ES 150', 'Lexus ES 150 4x4', 'Lexus ES 150', 'Lexus ES 140', 'Lexus ES 140 4x4', 'Lexus ES 140', 'Lexus ES 130', 'Lexus ES 130 4x4', 'Lexus ES 130', 'Lexus ES 120', 'Lexus ES 120 4x4', 'Lexus ES 120', 'Lexus ES 110', 'Lexus ES 110 4x4', 'Lexus ES 110', 'Lexus ES 100', 'Lexus ES 100 4x4', 'Lexus ES 100', 'Lexus ES 90', 'Lexus ES 90 4x4', 'Lexus ES 90', 'Lexus ES 80', 'Lexus ES 80 4x4', 'Lexus ES 80', 'Lexus ES 70', 'Lexus ES 70 4x4', 'Lexus ES 70', 'Lexus ES 60', 'Lexus ES 60 4x4', 'Lexus ES 60', 'Lexus ES 50', 'Lexus ES 50 4x4', 'Lexus ES 50', 'Lexus ES 40', 'Lexus ES 40 4x4', 'Lexus ES 40', 'Lexus ES 30', 'Lexus ES 30 4x4', 'Lexus ES 30', 'Lexus ES 20', 'Lexus ES 20 4x4', 'Lexus ES 20', 'Lexus ES 10', 'Lexus ES 10 4x4', 'Lexus ES 10', 'Lexus ES 0', 'Lexus ES 0 4x4', 'Lexus ES 0'.

Figura: E os veículos Lexus?



# Estudo de Caso

- Hipótese: potência é inversamente proporcional ao consumo
  - Pode ser confirmada de modo geral?

# Estudo de Caso

- Hipótese: potência é inversamente proporcional ao consumo
  - Pode ser confirmada de modo geral?

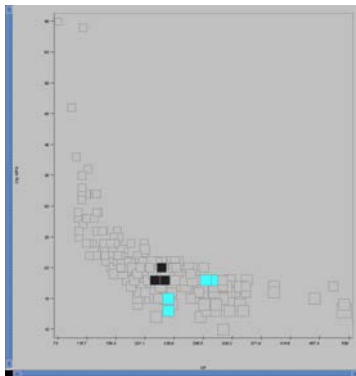


Figura: Scatterplot para veículos Lexus. Hipótese não se confirma.

- Considerando todo o conjunto de dados
  - As tendências identificadas nos sub-conjuntos se mantêm?
  - Existem instâncias com valores não preenchidos?
  - O que pode ser dito sobre os valores não preenchidos?
  - O que se pode afirmar sobre os dados como um todo (tendências, grupos, etc.)?

# Estudo de Caso

- Considerando todo o conjunto de dados
  - As tendências identificadas nos sub-conjuntos se mantêm?
  - Existem instâncias com valores não preenchidos?
  - O que pode ser dito sobre os valores não preenchidos?
  - O que se pode afirmar sobre os dados como um todo (tendências, grupos, etc.)?

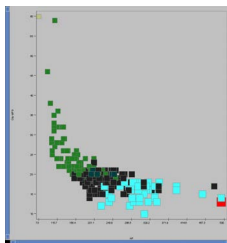


Figura: Scatterplot de todos os veículos.



# Estudo de Caso

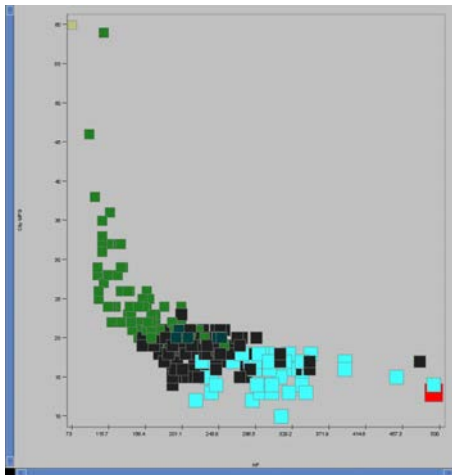


Figura: Scatterplot de todos os veículos.

# Sumário

- 1 Motivação
- 2 Introdução
  - O que é Visualização
  - História da Visualização
  - Visualização nos Dias Atuais
  - Relacionamento entre Visualização e outras áreas do conhecimento
  - O Processo de Visualização
- 3 Exemplo de análise
- 4 Papel do Usuário**
- 5 O Problema Fundamental
- 6 Sobre os slides do curso
- 7 Atividades
- 8 Referências

# Papel do Usuário

- A visualização pode ser executada com diferentes objetivos
  - **Exploração**: usuário quer entender, interpretar, levantar hipóteses sobre os dados
  - **Confirmação**: já existe uma hipótese(s), o usuário quer confirmá-la(s)
  - **Apresentação**: apresentar um conceito ou conjunto de fatos a um público



# Sumário

- 1 Motivação
- 2 Introdução
  - O que é Visualização
  - História da Visualização
  - Visualização nos Dias Atuais
  - Relacionamento entre Visualização e outras áreas do conhecimento
  - O Processo de Visualização
- 3 Exemplo de análise
- 4 Papel do Usuário
- 5 O Problema Fundamental**
- 6 Sobre os slides do curso
- 7 Atividades
- 8 Referências

## *Vis Design space*

- O espaço de possíveis soluções para exibir dados é enorme...
- Muitas opções, poucas soluções de fato adequadas
- Disciplina: conhecer um pouco sobre as opções de solução, e entender quais são adequadas para que situação

- 1 Motivação
- 2 Introdução
  - O que é Visualização
  - História da Visualização
  - Visualização nos Dias Atuais
  - Relacionamento entre Visualização e outras áreas do conhecimento
  - O Processo de Visualização
- 3 Exemplo de análise
- 4 Papel do Usuário
- 5 O Problema Fundamental**
- 6 Sobre os slides do curso
- 7 Atividades
- 8 Referências

- Limitações inerentes: computacionais, humanas, dos dispositivos de exibição
  - Computacionais: tempo de processamento e memória
  - Humanas: sistema perceptual, atenção, memória
  - Dispositivos: número finito de pixels, densidade de informação

# Sumário

- 1 Motivação
- 2 Introdução
  - O que é Visualização
  - História da Visualização
  - Visualização nos Dias Atuais
  - Relacionamento entre Visualização e outras áreas do conhecimento
  - O Processo de Visualização
- 3 Exemplo de análise
- 4 Papel do Usuário
- 5 O Problema Fundamental
- 6 Sobre os slides do curso**
- 7 Atividades
- 8 Referências

- fonte principal: livro Interactive Data Visualization, de Ward, Grinstein e Keim
- preparados a partir de material disponibilizado pelo Prof. Fernando Paulovich

# Sumário

- 1 Motivação
- 2 Introdução
  - O que é Visualização
  - História da Visualização
  - Visualização nos Dias Atuais
  - Relacionamento entre Visualização e outras áreas do conhecimento
  - O Processo de Visualização
- 3 Exemplo de análise
- 4 Papel do Usuário
- 5 O Problema Fundamental
- 6 Sobre os slides do curso
- 7 Atividades**
- 8 Referências

# Atividade 1 (grad)

- Empregar uma ferramenta de visualização (Excel, weave, xmdvtool, etc.) para analisar algum conjunto de dados disponível (site do livro)
  - 1 Leia o conjunto de dados inteiro no programa
  - 2 Selecione um subconjunto dos dados que contenha alguma correlação óbvia (visualização exploratória)
  - 3 Estabeleça uma hipótese e verifique-a no conjunto de dados como um todo (visualização confirmatória)
  - 4 Faça uma apresentação slides de suas conclusões (visualização para apresentação)
  - 5 Submeta arquivo com análise no Moodle (SCC0252) (Atividade 1)
  - 6 arquivo <Ativ1-seuNúmeroUSP>.pdf (max. 5 slides)



## Atividade 1 (pós)

- Leitura do artigo "Why A Diagram Is (Sometimes) Worth Ten Thousand Words (Larkin & Simon 1987)"
- Submeta arquivo com seu resumo dos principais pontos no Moodle (SCC5836) (Atividade 1)
- arquivo <Ativ1-seuNúmeroUSP>.pdf (2 páginas texto máx.)

# Sugestão de leitura (todos)

- Sobre a Teoria do Gestalt

- <http://graphicdesign.spokanefalls.edu/tutorials/process/gestaltprinciples/gestaltprinc.htm>
- <https://www.smashingmagazine.com/2014/03/design-principles-visual-perception-and-the-principles-of-gestalt/>

# Sumário

- 1 Motivação
- 2 Introdução
  - O que é Visualização
  - História da Visualização
  - Visualização nos Dias Atuais
  - Relacionamento entre Visualização e outras áreas do conhecimento
  - O Processo de Visualização
- 3 Exemplo de análise
- 4 Papel do Usuário
- 5 O Problema Fundamental
- 6 Sobre os slides do curso
- 7 Atividades
- 8 Referências

# Referências

- **[Munzner, 2015]** Tamara Munzner, Visualization Analysis & Design, CRC Press.
- **[Lyman & Varian, 2003]** Peter Lyman and Hal R. Varian, How Much Information, 2003; [www2.sims.berkeley.edu/research/projects/how-much-info/](http://www2.sims.berkeley.edu/research/projects/how-much-info/)
- **[Tanaka, 1998]** Jennifer Tanaka, Drowning in Data, Newsweek, 4/28/98, p. 85
- **[Cetron & Davies, 1991]** Marvin Cetron and Owen Davies, Crystal Globe, New York, St. Martin's Press, 1991, pp. 361-2
- **[Cetrin & Davies, 1989]** Marvin Cetron and Owen Davies, American Renaissance, New York, St. Martin's Press, 1989, p. 65
- **[Gladwell, 2008]** Malcolm Gladwell, Fora de Série: Outliers, Sextante, 2008.