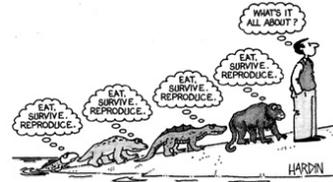


Apresentação Gráfica de Informação Quantitativa

Renato Vicente
GRIFE-EACH-USP/2008

Perguntas

1. Para que precisamos de gráficos?
2. Como surgiu a idéia?
3. Quais são seus elementos básicos?
4. Poderia dar exemplos?
5. Como posso confeccioná-los?
6. Como produzo gráficos com qualidade?
7. Resumindo



Para que servem os gráficos?

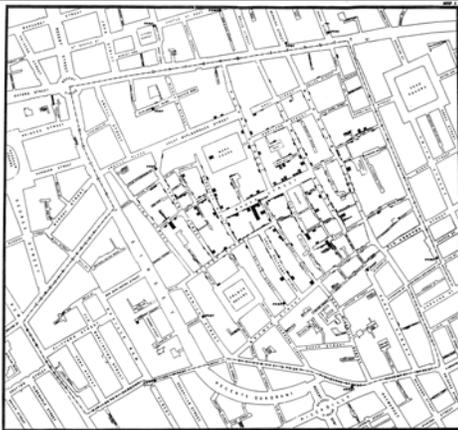
Epidemia de Cólera em Londres (set/1854)



John Snow



Análise de John Snow para a
epidemia de Cólera em Londres
(set/1854)



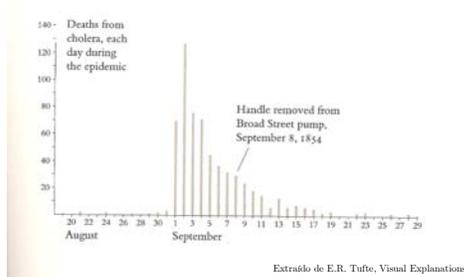
Análise de John Snow para a epidemia de Cólera em Londres

BROADWICK
STREET W1
CITY OF WESTMINSTER



Hoje é possível visitar a bomba d'água que serviu como evidência para a conexão causal entre o cólera e a contaminação da água. Há um pub em homenagem a John Snow.

Análise de John Snow para a epidemia de Cólera em Londres (set/1854)



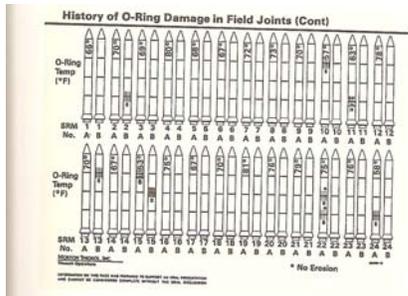
A epidemia já estava recuando quando a bomba de água foi desativada.

Ônibus espacial Challenger (28/jan/1986)



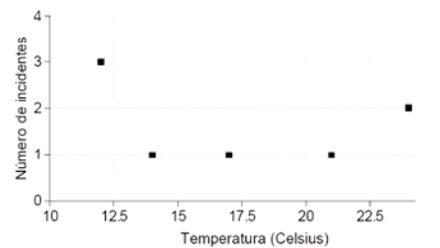
O ônibus espacial Challenger explodiu em 1986 devido a uma falha em um anel de vedação em um dos foguetes a propulente sólido

Ônibus espacial Challenger (28/jan/1986)



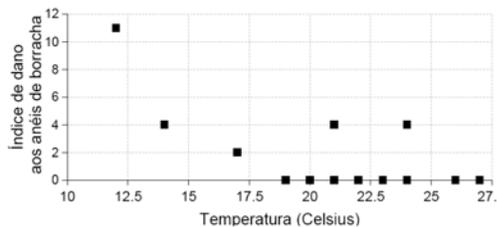
Eventos em testes com foguetes sólidos da Morton Thiokol Inc.

Explosão do ônibus espacial Challenger (28/jan/1986)



NÚMERO DE INCIDENTES COMO FUNÇÃO DA TEMPERATURA. Somente são exibidos lançamentos com ocorrência de incidentes. A previsão de temperatura para o dia do lançamento era algo em torno de 0 °C. Você autorizaria o lançamento? Por que?

Explosão do ônibus espacial Challenger (28/jan/1986)



HISTÓRICO DE DANOS AOS ANÉIS DE VEDAÇÃO COMO FUNÇÃO DA TEMPERATURA. A previsão de temperatura para o dia do lançamento era algo em torno de 0 °C. Você autorizaria o lançamento? Por que?

Explosão do ônibus espacial Challenger (28/jan/1986)



I took this stuff that I got out of your seal and I put it in ice water, and I discovered that when you put some pressure on it for a while and then undo it it doesn't stretch back. It stays the same dimension. In other words, for a few seconds at least and more seconds than that, there is no resilience in this particular material when it is at a temperature of 32 degrees.

I believe that has some significance to our problem. Richard P. Feynman



Como surgiu a idéia?

Representações Egípcias 2000 a.C.



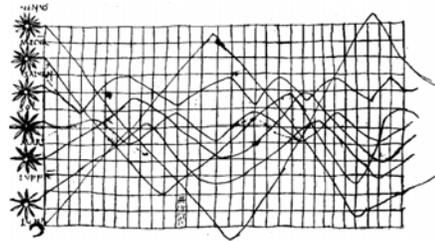
As decorações egípcias eram desenhadas como diagramas codificados para comunicar com clareza eventos do passado.

Representações Egípcias 1400 a.C.



"O Jardim de Nebamun" (1400 aC), Tebas, atualmente no Museu Britânico. Há aqui uma preocupação em representar diagrammaticamente, de forma identificável, as espécies presentes no jardim.

Movimento de Planetas: Século 10 (ou 11)



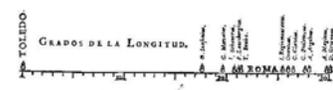
A representação acima é o mais antigo registro gráfico dos movimentos planetários encontrada em um texto de instrução para monastérios do século 10 (ou 11).

Nicole d'Oresme (1323-1382)

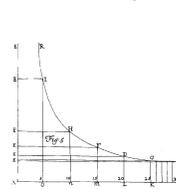


Primeira proposta de representação gráfica de dependência entre valores quantitativos de variáveis publicado em manuscrito de Nicole d'Oresme de 1350.

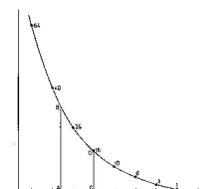
Século de Newton (Séc 17)



Medidas de longitude de Michael van Langren (1644)



Função contínua de Halley



Espectativa de vida de Christiaan Huygens (1669)

Quais são os elementos básicos de um gráfico?

Terminologia: Dois conjuntos superpostos de dados

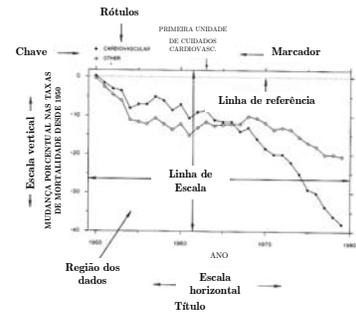


Figura 1. TAXAS DE MORTALIDADE AJUSTADAS PELA IDADE. Os dados representam o percentual de mudança nas taxas de mortalidade devido a problemas cardiovasculares e a outras doenças nos EUA

Terminologia: Dois conjuntos justapostos de dados

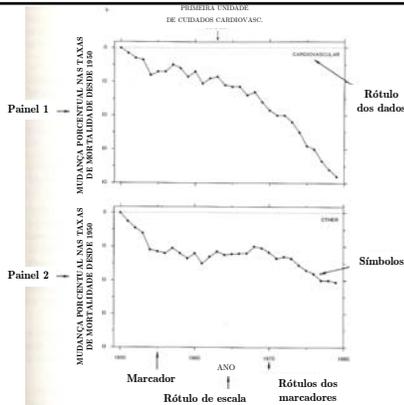
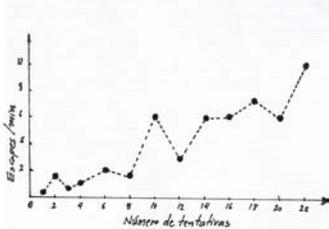


Figura 2. TAXAS DE MORTALIDADE AJUSTADAS PELA IDADE. Os dados representam o percentual de mudança nas taxas de mortalidade devido a problemas cardiovasculares e a outras doenças nos EUA

Como confeccioná-los ?

Ferramentas de Confeção

1. Papel milimetrado



APRENDIZADO EM GATOS. Curva de aprendizagem de gatos expostos à tarefa de escapar de um labirinto. O gráfico representa a rapidez da execução da tarefa (em escapes por minuto) contra o número de tentativas.

Ferramentas de Confeção

2. Planilhas Eletrônicas

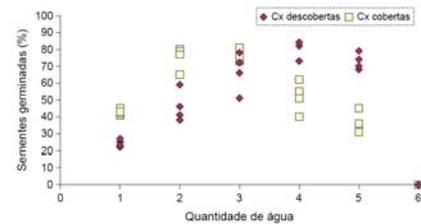
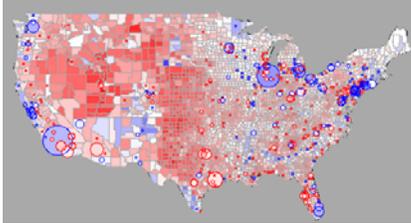


Figura 1. EFEITO DA QUANTIDADE DE ÁGUA SOBRE A GERMINAÇÃO DE SEMENTES. Os dados correspondem aos percentuais de sementes germinadas após duas semanas em quatro caixas idênticas. O experimento foi repetido para 6 quantidades diferentes de água e para caixas descobertas para retardar a evaporação e cobertas.

Ferramentas de Confeção

3. Software Estatístico R

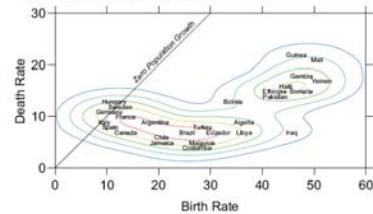


RESULTADO DAS ELEIÇÕES PRESIDENCIAIS DE 2004 NOS EUA.
A área dos círculos é representa o tamanho da diferença de votos. Áreas azuis representam vitória de Kerry e em vermelho vitória de Bush. Kerry obteve grandes margens em grandes cidades nas costas enquanto Bush venceu em pequenas cidades do meio oeste.

Ferramentas de Confeção

4. Gramática dos Gráficos (implementada em R)

```
ELEMENT: point(position(birth*death), size(0), label(country))
ELEMENT: contour(position(
  smooth.density.kernel(Lepanochnikov,join(birth*death)),
  color.hue()))
GUIDE: form.line(position(0,0),30,30), label("Zero Population Growth")
GUIDE: axis(dim(1), label("Birth Rate"))
GUIDE: axis(dim(2), label("Death Rate"))
```

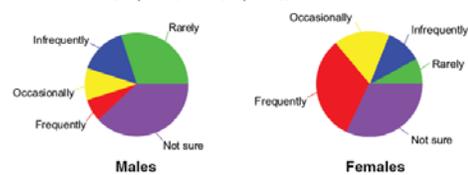


TAXAS DE MORTALIDADE E NATALIDADE POR 100.000 HABITANTES

Poderia dar exemplos ?

Gráfico de Setores

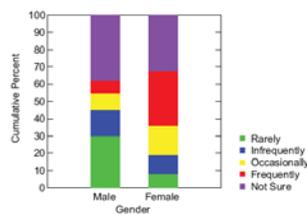
```
DATA: response = Response
DATA: gender = Gender
SCALE: cat(dim(1),
  values("Rarely", "Infrequently", "Occasionally", "Frequently", "Not Sure"))
SCALE: cat(dim(2), values("Female", "Male"))
COORD: rect(dim(2), polar.theta(dim(1)))
ELEMENT: interval.stack(position(summary.proportion(response*gender)),
  label(response), color(response))
```



PERCEÇÃO DE VIÉS DE GÊNERO NA AVALIAÇÃO DE ARTIGOS SUBMETIDOS À REVISTAS CIENTÍFICAS

Gráfico de Barras

```
ELEMENT: interval.stack(position(summary.proportion(gender*response),
  color(response)))
```



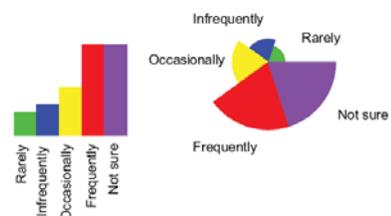
PERCEÇÃO DE VIÉS DE GÊNERO NA AVALIAÇÃO DE ARTIGOS SUBMETIDOS À REVISTAS CIENTÍFICAS

Barras e Rosa de Nightingale

```
ELEMENT: interval(position(response), color(response), size(2))
```

```
COORD: polar()
```

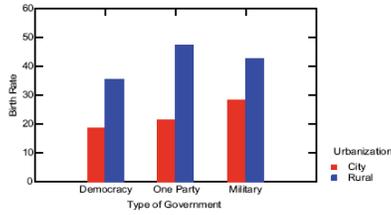
```
ELEMENT: interval(position(response), color(response), size(2))
```



PERCEÇÃO DE VIÉS DE GÊNERO NA AVALIAÇÃO DE ARTIGOS SUBMETIDOS À REVISTAS CIENTÍFICAS

Gráfico de Barras 2

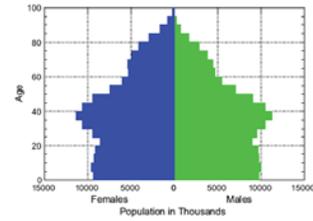
ELEMENT: `interval.dodge(position(summary.mean(gov2birth)), color(urban))`



TAXA DE NATALIDADE SEGUNDO TIPO DE GOVERNO E LOCAL DE RESIDÊNCIA

Gráfico de Barras 3

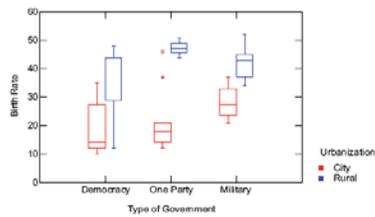
COORD: `transpose(mirror(dim(3), rect(dim(1, 2))))`
 ELEMENT: `interval(position(age2pop2sex), color(sex), size(2))`



COMPOSIÇÃO DA POPULAÇÃO DOS EUA EM 1980 POR GÊNERO

Box Plot

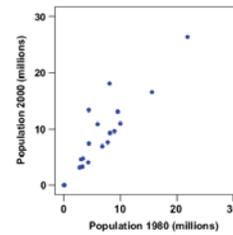
ELEMENT: `schema.dodge(position(bin.quantile.letter(gov2birth)), color(urban))`



TAXA DE NATALIDADE SEGUNDO TIPO DE GOVERNO E LOCAL DE RESIDÊNCIA

Dispersão

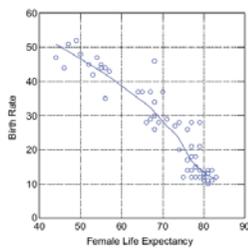
ELEMENT: `point(position(pop19802pop2000))`



CRESCIMENTO POPULACIONAL DE UM CONJUNTO DE CIDADES.

Dispersão com LOESS

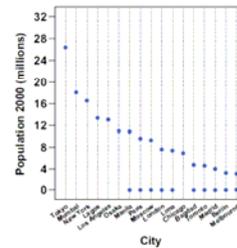
ELEMENT: `line(position(smooth.loess(female2birth)))`
 ELEMENT: `point(position(female2birth))`



TAXA DE NATALIDADE VERSUS EXPECTATIVA DE VIDA FEMININA COM ALISAMENTO LOESS PARA AUXILIAR A VISUALIZAÇÃO

Gráfico de Pontos (Dot Plot)

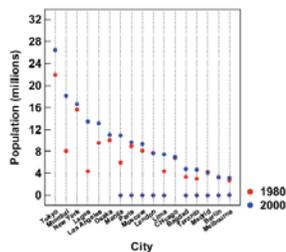
ELEMENT: `point(position(city2pop2000))`



POPULAÇÃO DE UM CONJUNTO DE CIDADES.

Gráfico de Pontos Com Sobreposição (Dot Plot)

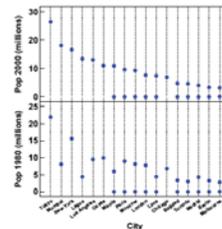
DATA: p1980 = "1980"
 DATA: p2000 = "2000"
 ELEMENT: point(position(city*(pop1980+pop2000)), color(p1980 + p2000))



COMPARAÇÃO DE POPULAÇÕES PARA UM CONJUNTO DE CIDADES

Gráfico de Pontos Com Justaposição (Dot Plot)

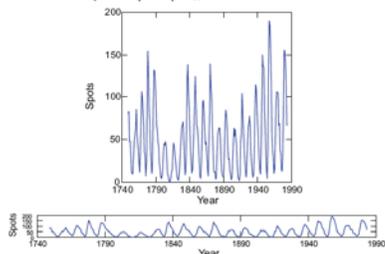
DATA: p1980 = "Pop 1980 (millions)"
 DATA: p2000 = "Pop 2000 (millions)"
 ELEMENT: point(position(city*(pop1980/p1980+pop2000/p2000)))



COMPARAÇÃO DE POPULAÇÕES PARA UM CONJUNTO DE CIDADES

Série Temporal

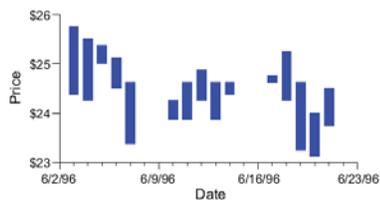
ELEMENT: line(position(year*spots))
 COORD: stretch(factor(2.0, 0.111))
 ELEMENT: line(position(year*spots))



MANCHAS SOLARES

Série Temporal com Intervalos

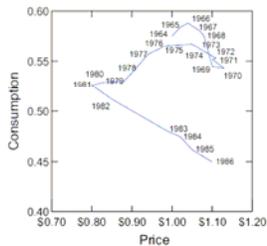
SCALE: time(dim(1))
 GUIDE: axis(dim(1), format("mm/dd/yy"))
 ELEMENT: interval(position(region.spread.range(date*(high+low))))



VARIAÇÃO DE PREÇOS NA BOLSA DE VALORES

Trajetória

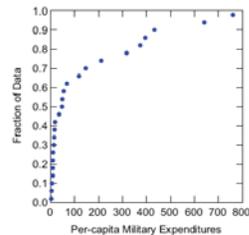
ELEMENT: path(position(price*consumption), label(year))



EVOLUÇÃO TEMPORAL DE PREÇOS E CONSUMO DE CIGARROS.

Gráfico de Quantis

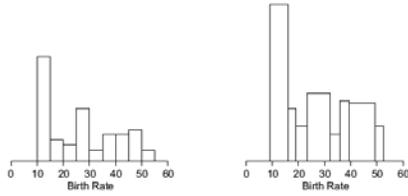
TRANS: rml = prank(military)
 ELEMENT: point(position(military*rml))



DISTRIBUIÇÃO DE GASTOS MILITARES

Histograma 1

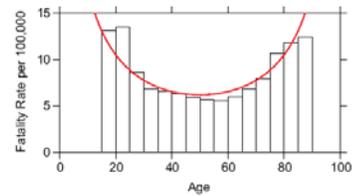
ELEMENT: `interval(position(summary.count(bin.rect(birth))))`
 ELEMENT: `interval(position(summary.count(bin.voronoi(birth))))`



HISTOGRAMA DE TAXAS DE NATALIDADE COM DOIS CRITÉRIOS DE SEGMENTAÇÃO DE ESCALA (REGULAR E VORONOI)

Histograma 2

DATA: `weight = col(source("traffic.txt"), name("percent"), weight())`
 ELEMENT: `interval(position(summary.sum(bin.rect(age*weight), dim(1))))`
 ELEMENT: `line(position(age*arcsin))`

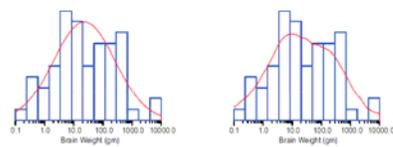


FREQUENCIA DE MORTES NO TRANSITO POR IDADE

Histograma 3

SCALE: `log(dim 1), base(10)`
 ELEMENT: `interval(position(summary.proportion(bin.rect(brainweight)), color(color.blue)))`
 ELEMENT: `line(position(smooth.density.normal(brainweight)), color(color.red))`

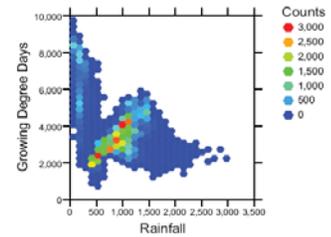
SCALE: `log(dim 1), base(10)`
 ELEMENT: `interval(position(summary.proportion(bin.rect(brainweight)), color(color.blue)))`
 ELEMENT: `line(position(smooth.density.kernel(brainweight)), color(color.red))`



DISTRIBUIÇÃO DE MASSA CEREBRAL PARA VÁRIAS ESPÉCIES COM ESCALA LOGARITMICA E INFERÊNCIA DE DISTRIBUIÇÃO

Histograma em 2D

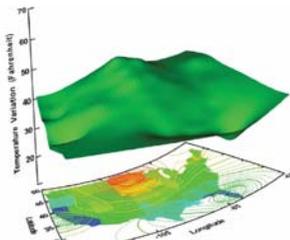
ELEMENT: `polygon(position(bin.hex(rainfall*degdays)), color.hue(summary.count()))`



DISTRIBUIÇÃO DE DIAS POR VARIAÇÃO DE TEMPERATURA E PRECIPITAÇÃO DE CHUVA

Gráficos Tridimensionais

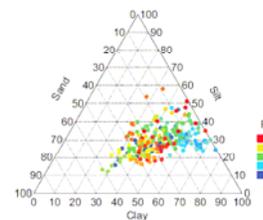
DATA: `longitude, latitude = map(source("US States"))`
 TRANS: `sw = diff(summer, winter)`
 COORD: `rect(dim(1, 2, 3), project.stereodim(1, 2))`
 ELEMENT: `polygone(position(longitude*latitude), color.hue(sw))`
 ELEMENT: `contour(position(smooth.mean.cauchy(lon*lat*sw)), color.hue())`
 ELEMENT: `surface(position(smooth.mean.cauchy(lon*lat*sw)))`



ALISAMENTO DA DISTRIBUIÇÃO DE TEMPERATURAS NOS EUA

Coordenadas Triangulares

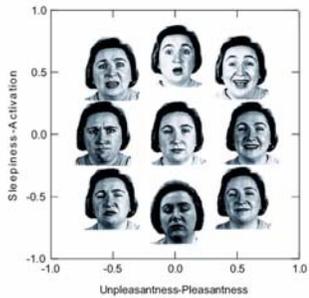
COORD: `triangular(dim(1, 2, 3, 4))`
 ELEMENT: `point(position(sand*silt*clay), color.hue(plot))`



COMPOSIÇÃO DE AMOSTRAS DE SOLO

Faces

DATA: face = *link*("faces")
ELEMENT: *point*(*position***pleasantness***activation*), *shape*.*image*(*face*)



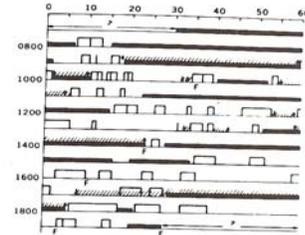
Intensidade de emoções

Como produzo gráficos com qualidade ?

Princípios de construção de gráficos

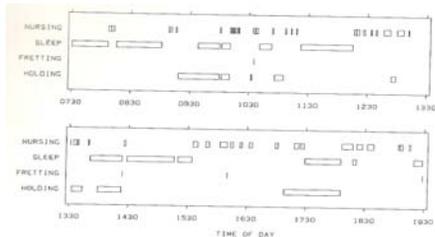
1. Visualização clara
2. Compreensão clara
3. Escolha apropriada de escalas

Visualização clara: Os dados devem se destacar em relação aos recursos gráficos.



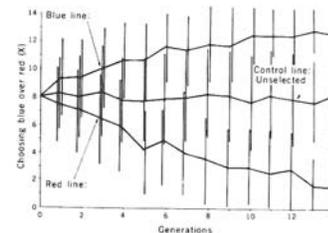
ATIVIDADES DE MULHERES KUNG E SEUS BEBÊS. Os eixos vertical e horizontal indicam as horas e os minutos, respectivamente. Retângulos claros e barras verticais indicam amamentação, a letra F indica choro do bebê, as barras escuras indicam sono do bebê e os achurados indicam mãe com bebê no colo. O gráfico é confuso.

Visualização clara: Os dados devem se destacar em relação aos recursos gráficos.



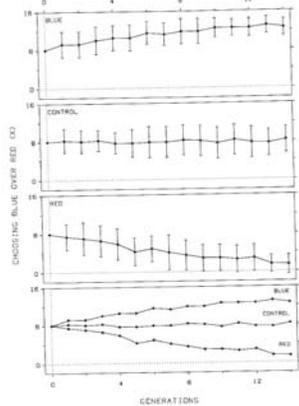
ATIVIDADES DE MULHERES KUNG E SEUS BEBÊS. As barras indicam os momentos dedicados a cada uma das seguintes atividades: amamentação (nursing), sono (sleep), choro (fretting) e colo (holding). Sem necessidade de elementos gráficos supérfluos, os dados se destacam nessa versão.

Visualização clara: Os dados devem se destacar em relação aos recursos gráficos.



EVOLUÇÃO DE PREFERÊNCIAS POR CORES EM GALINÁCEOS. O experimento consiste na seleção de gerações de animais segundo preferências pela cor vermelha ou azul. No centro há uma linha representando o controle sem seleção. As barras indicam a distribuição de preferências. As barras tornam o gráfico superposto confuso.

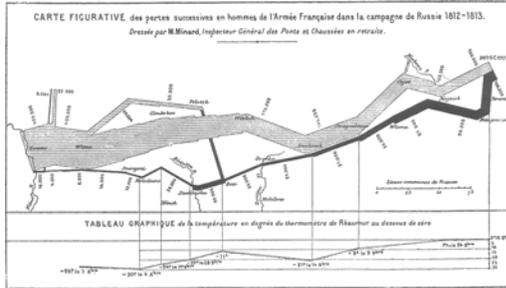
EVOLUÇÃO DE
PREFERÊNCIAS POR
CORES EM GALINÁCEOS.
O experimento consiste na
seleção de gerações de animais
segundo preferências pela cor
vermelha ou azul. As barras
indicam a distribuição de
preferências. A justaposição é
a escolha mais adequada.



Compreensão clara

1. Descreva na legenda tudo o que estiver representado no gráfico.
2. Dê particular atenção às características importantes do gráfico.
3. Descreva as conclusões extraídas do gráfico.

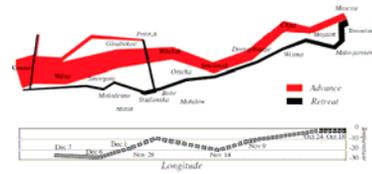
Excelência Gráfica



CAMPANHA DE NAPOLEÃO CONTRA A RÚSSIA. O gráfico mostra a marcha das tropas napoleônicas contra Moscou em 1812-1813. A largura da figura representa o número de soldados. A curva cinzenta representa a marcha de ida e a preta indica a marcha de volta. Abaixo há um gráfico indicando as temperaturas durante a marcha de volta. As tropas pereceram no rigoroso inverno russo.

Campanha Napoleônica

```
TRANS: date = m/s/d date, font, "linear")
TRANS: ldate = left(date, -1)
TRANS: days = diff(date, ldate)
GRAPH: begin origin(0, 0), scale(12cm, 4cm)
ELEMENT: point position=long=latc, label(city), size(0)
ELEMENT: path position=long=latc, size(our-svovs), color(direction),
split(group)
GUIDE: legend, color(dim 1)
GRAPH: end
GRAPH: begin origin(0, -2cm), scale(12cm, 2cm)
ELEMENT: path position=long=temp, label(date), texture, grain=ativity(days)
color(brightness=5)
GUIDE: axis(dim 1), label("Longitude")
GUIDE: axis(dim 2), label("Temperature")
GRAPH: end
```



Excelência Gráfica

1. A excelência gráfica é a apresentação bem projetada de dados interessantes, uma combinação entre **substância, estatística e design gráfico**.
2. A excelência gráfica consiste da comunicação de idéias com **clareza, precisão e eficiência**.
3. A excelência gráfica é o que dá ao leitor o **maior número de idéias no menor tempo com o menor uso de tinta de no menor espaço**.
4. A excelência gráfica requer que se conte a **verdade dobre os dados**.

Edward. R. Tufte

Resumindo

1. Para que precisamos de gráficos?
2. Como surgiu a idéia?
3. Quais são seus elementos básicos?
4. Poderia dar exemplos?
5. Como posso confeccioná-los?
6. Como produzo gráficos com qualidade?
7. Resumindo

