


INSTITUTO FEDERAL
São Paulo
Câmpus Campinas

Visual eficaz

D1STO - Storytelling com Dados

Especialização em Ciência de
Dados





***“Se uma imagem vale mais do
que mil palavras, então diga
isto com uma imagem.”***

Millôr Fernandes

91

Texto simples

| | A | B | C |
|------------|-----|-----|-----|
| Category 1 | 15% | 22% | 42% |
| Category 2 | 40% | 36% | 20% |
| Category 3 | 35% | 17% | 34% |
| Category 4 | 30% | 29% | 26% |
| Category 5 | 55% | 30% | 58% |
| Category 6 | 11% | 25% | 49% |

Tabela

| | A | B | C |
|------------|-----|-----|-----|
| Category 1 | 15% | 22% | 42% |
| Category 2 | 40% | 36% | 20% |
| Category 3 | 35% | 17% | 34% |
| Category 4 | 30% | 29% | 26% |
| Category 5 | 55% | 30% | 58% |
| Category 6 | 11% | 25% | 49% |

Mapa de calor

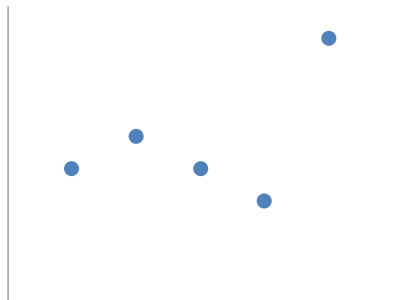
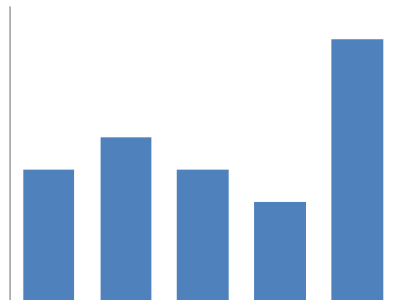
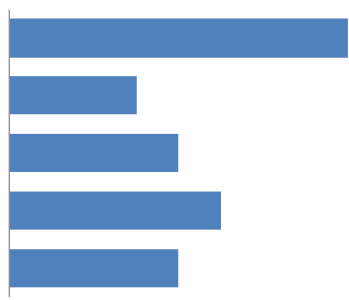


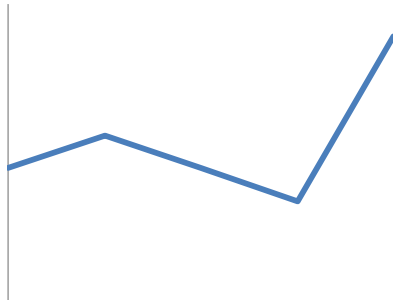
Gráfico de dispersão



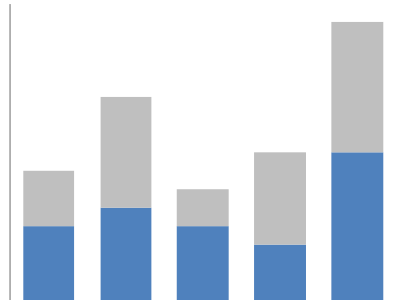
Barra vertical



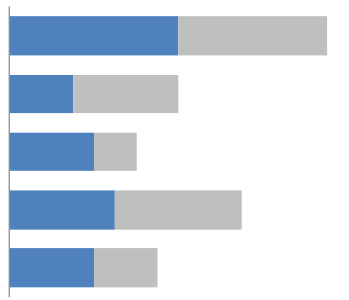
Barra horizontal



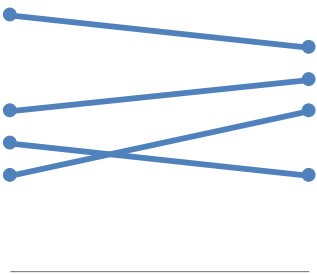
Linha



Barras verticais empilhadas



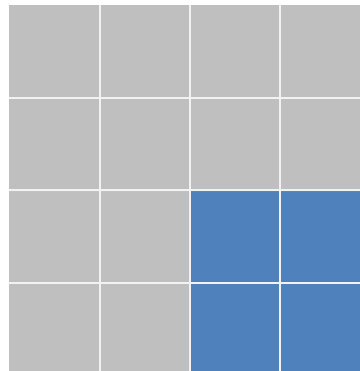
Barras horizontais empilhadas



Mapa de inclinação



Cascata



Área quadrada

Texto simples

Útil para quando se tem **1 ou 2 números** para mostrar

Caso contrário, perde-se a força se utilizado num **gráfico** ou **tabela**

Também desmitifica que gráficos e tabelas são os únicos meios de comunicar os dados

Percepção visual (20 é menos da metade de 41) pode ficar comprometida

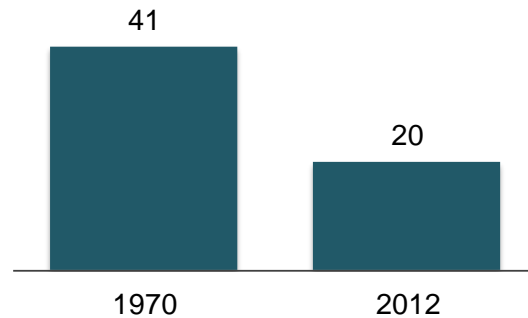
Algo também que pode ser feito é dizer (com cautela, para não perder o contexto):

“O número de filhos com mães tradicionais que ficam em casa diminuiu mais de 50% entre 1970 e 2012”

FIG0202

Children with a "Traditional" Stay-at-Home Mother

% of children with a married stay-at-home mother with a working husband



Note: Based on children younger than 18. Their mothers are categorized based on employment status in 1970 and 2012.

Source: Pew Research Center analysis of March Current Population Surveys Integrated Public Use Microdata Series (IPUMS-CPS), 1971 and 2013

Adapted from PEW RESEARCH CENTER

FIG0203

20%

of children had a **traditional stay-at-home mom** in 2012, compared to 41% in 1970

Tabelas

Interagem com nosso **sistema verbal**. Nós as lemos!

Útil para comunicar várias medidas diferentes

Dados **sempre** em primeiro plano; design é secundário

Dica

Raramente se usa tabelas em apresentações ao vivo. O público tende a ler os dados, e não prestar a atenção no que o apresentador fala.

FIG0204

Heavy borders

| Group | Metric A | Metric B | Metric C |
|---------|----------|----------|----------|
| Group 1 | \$X.X | Y% | Z,ZZZ |
| Group 2 | \$X.X | Y% | Z,ZZZ |
| Group 3 | \$X.X | Y% | Z,ZZZ |
| Group 4 | \$X.X | Y% | Z,ZZZ |
| Group 5 | \$X.X | Y% | Z,ZZZ |

Light borders

| Group | Metric A | Metric B | Metric C |
|---------|----------|----------|----------|
| Group 1 | \$X.X | Y% | Z,ZZZ |
| Group 2 | \$X.X | Y% | Z,ZZZ |
| Group 3 | \$X.X | Y% | Z,ZZZ |
| Group 4 | \$X.X | Y% | Z,ZZZ |
| Group 5 | \$X.X | Y% | Z,ZZZ |

Minimal borders

| Group | Metric A | Metric B | Metric C |
|---------|----------|----------|----------|
| Group 1 | \$X.X | Y% | Z,ZZZ |
| Group 2 | \$X.X | Y% | Z,ZZZ |
| Group 3 | \$X.X | Y% | Z,ZZZ |
| Group 4 | \$X.X | Y% | Z,ZZZ |
| Group 5 | \$X.X | Y% | Z,ZZZ |

Mapa de calor

Une detalhes de uma tabela com **pistas visuais**

Ajuda a “cortar caminho” na leitura dos dados

Saturação de cores leva aos pontos de interesse

Aplicativos (p. ex. Excel, Data Studio) têm esse recurso nativo

FIG0205

Table

| | A | B | C |
|------------|-----|-----|-----|
| Category 1 | 15% | 22% | 42% |
| Category 2 | 40% | 36% | 20% |
| Category 3 | 35% | 17% | 34% |
| Category 4 | 30% | 29% | 26% |
| Category 5 | 55% | 30% | 58% |
| Category 6 | 11% | 25% | 49% |

Heatmap

LOW-HIGH

| | A | B | C |
|------------|-----|-----|-----|
| Category 1 | 15% | 22% | 42% |
| Category 2 | 40% | 36% | 20% |
| Category 3 | 35% | 17% | 34% |
| Category 4 | 30% | 29% | 26% |
| Category 5 | 55% | 30% | 58% |
| Category 6 | 11% | 25% | 49% |

Gráficos

Tabelas interagem com nosso **sistema verbal**

Gráficos interagem com nosso **sistema visual**

Normalmente, caem em 4 categorias:

PONTOS • LINHAS • BARRAS • ÁREA

Pontos

Gráfico de dispersão

Útil para mostrar **relação** entre 2 coisas

Codifica dados simultaneamente (eixos x e y) para exprimir a relação

Mais usados no campo científico

Por vezes, complicado para públicos do mundo dos negócios

Por exemplo, digamos que gerenciamos uma frota de ônibus e queremos entender a relação entre as milhas dirigidas e o custo por milha.

Gráfico de dispersão ligeiramente modificado, projetado para levar os olhos mais rapidamente para custos acima da média (abaixo de cerca de 1.700 milhas ou acima de 3.300 milhas).

FIG0206

Cost per mile by miles driven

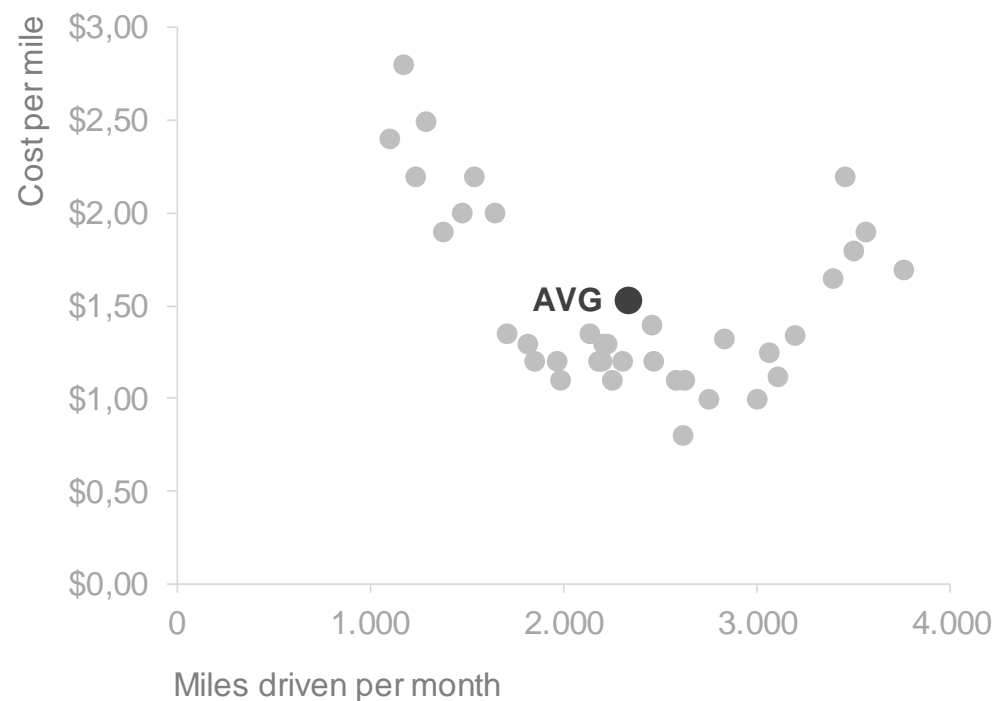
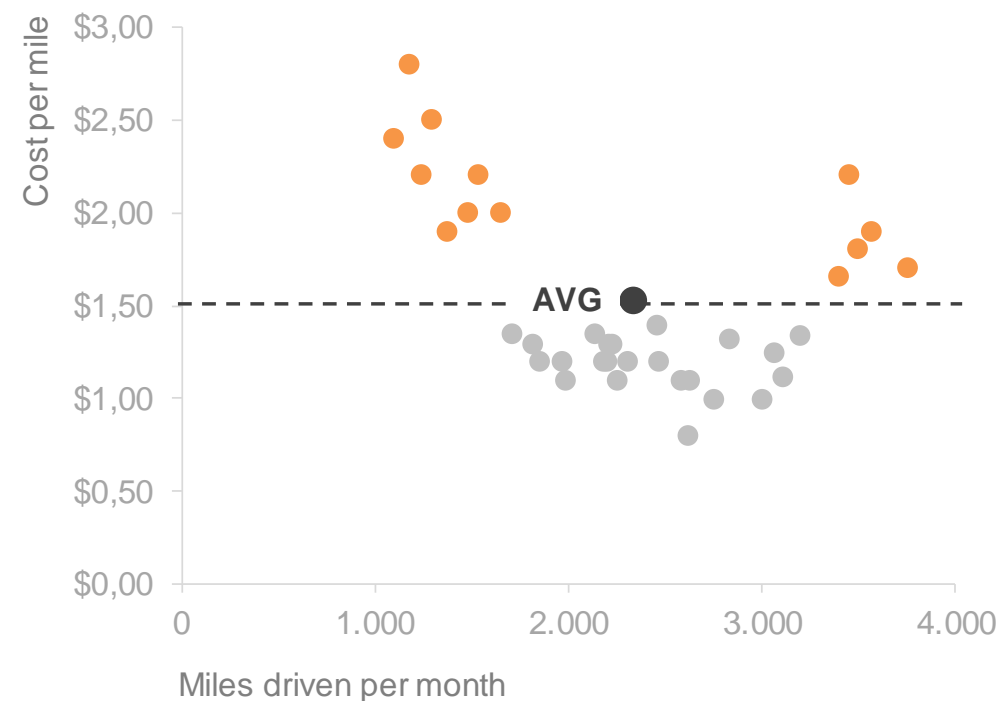


FIG0207

Cost per mile by miles driven



Linhas

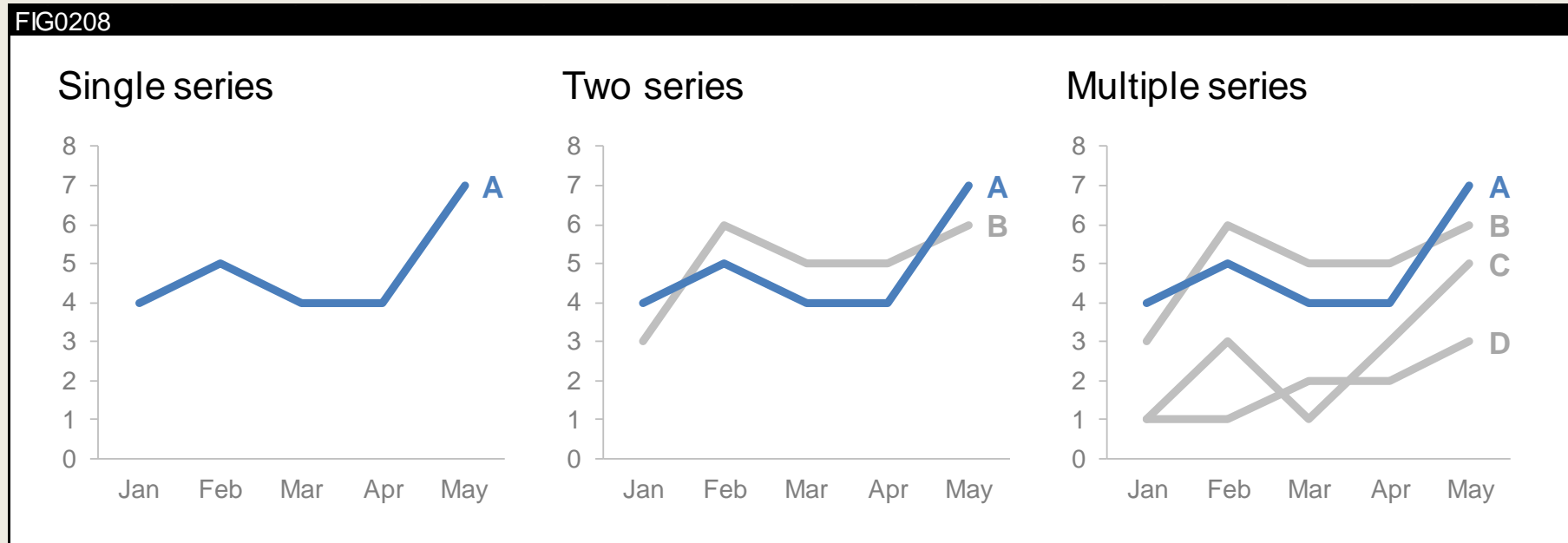
Gráfico de linhas

Mais usados para registrar **dados contínuos**

Indica uma conexão (p. ex. tempo: dias, meses, etc.)

Pode não fazer sentido para dados categóricos

Eixo x (horizontal) deve estar em **intervalos coerentes**



EXEMPLO DE GRÁFICO DE LINHAS COMBINADO

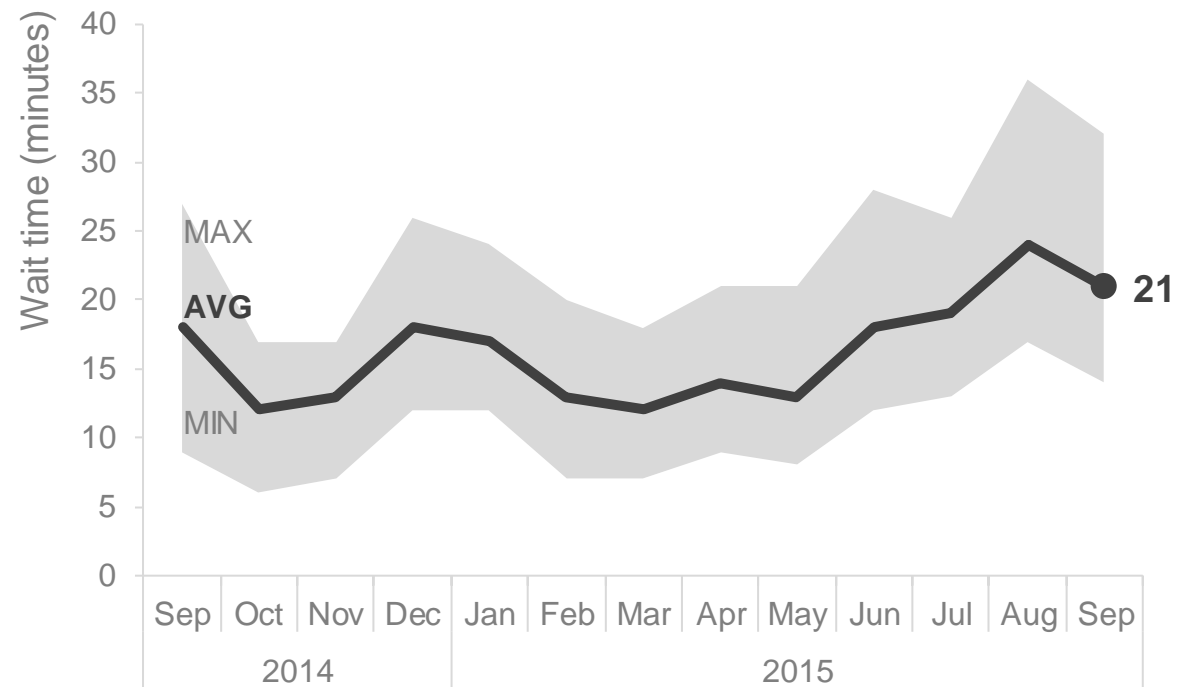
Linhas (séries) podem representar

- Resumo estatístico
- Intervalo p. ex.:
 - Nível de confiança
 - Mínimo e máximo

FIG0209

Passport control wait time

Past 13 months



Linhas

Gráfico de inclinação

Úteis quando se tem **2 períodos** de tempo ou pontos de **comparação**

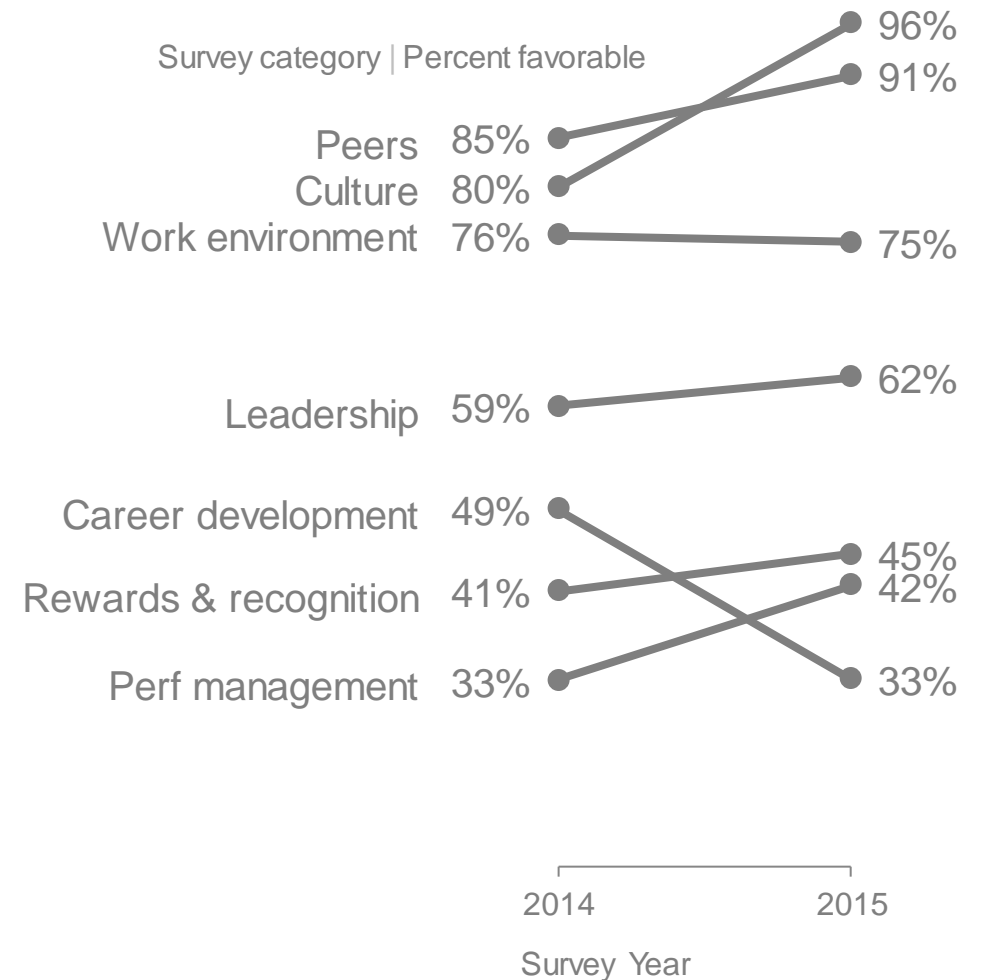
Mostra rapidamente aumentos/diminuições, por meio de **valores** e **visual**

Cuidado quando muitas linhas estão se sobrepondo

Requer personalização no Excel. Tutorial em <https://exceloffthegrid.com/the-ultimate-guide-to-slopegraphs-in-excel/>

FIG0210

Employee feedback over time

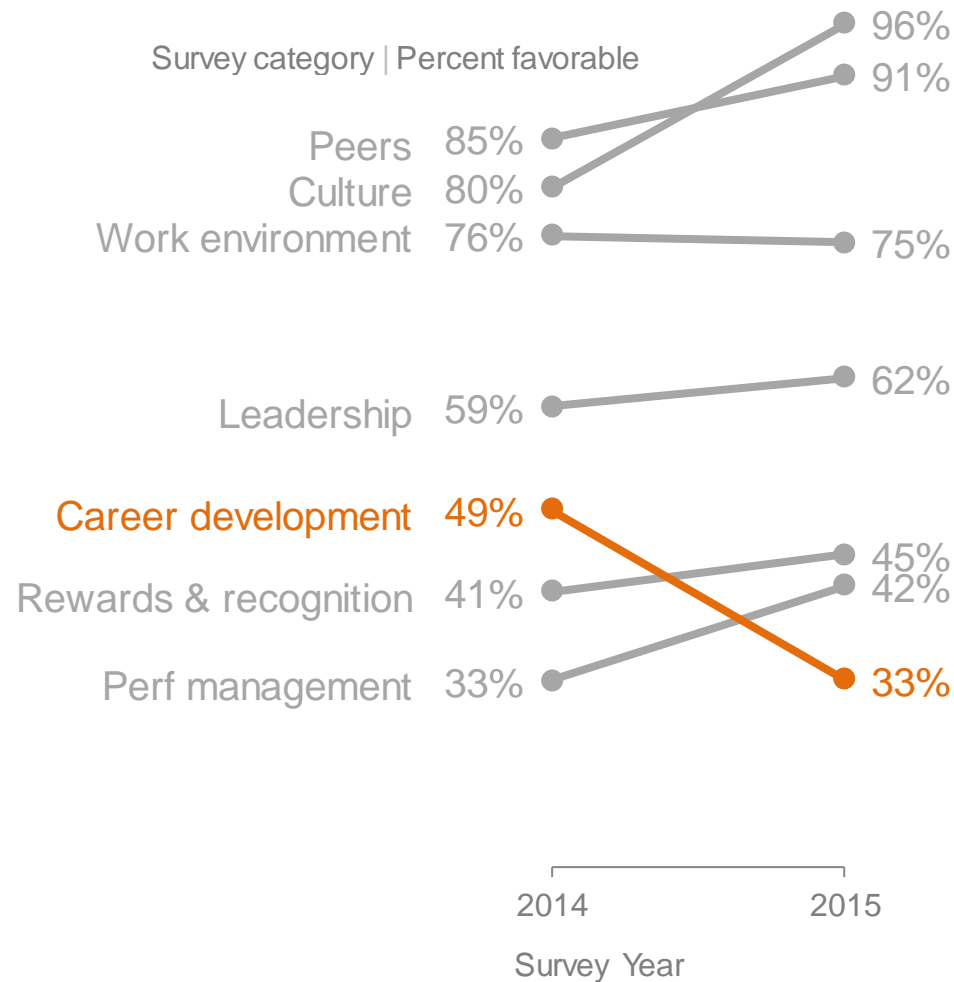


EXEMPLO DE GRÁFICO DE INCLINAÇÃO COM ATRIBUTO PRÉ-ATENTIVO

Enfatizando uma série por vez:
podemos chamar a atenção para a categoria do exemplo anterior que **diminuiu ao longo do tempo.**

FIG0211

Employee feedback over time

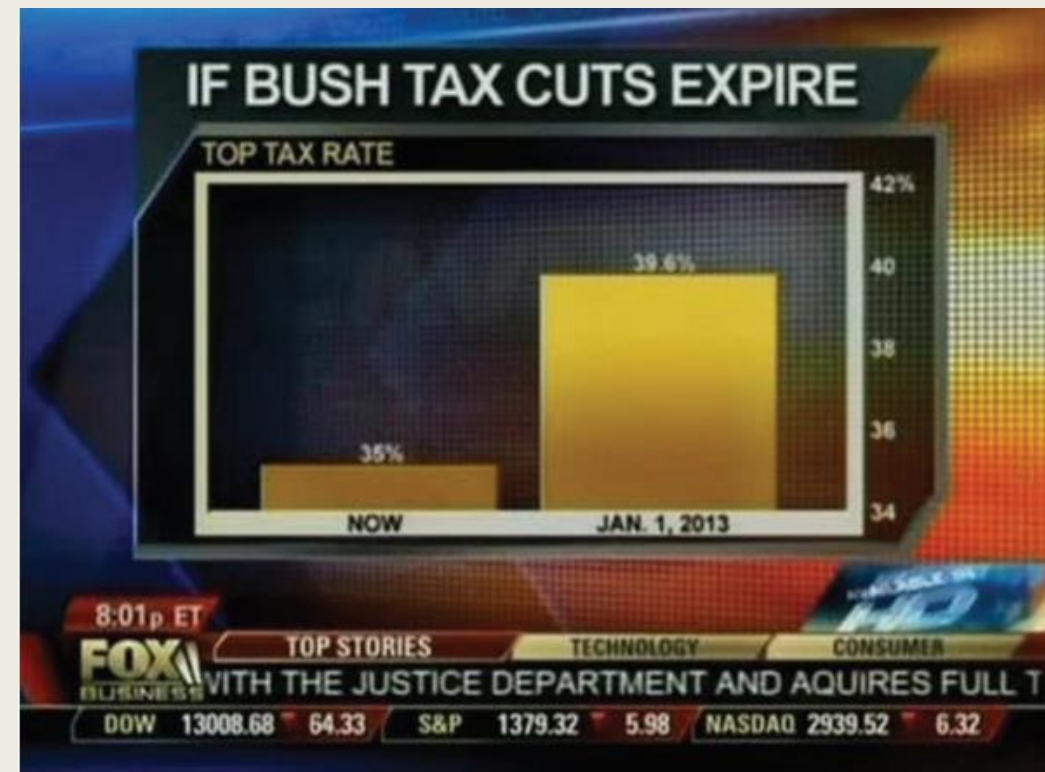


Barras

Muito comuns. Todo mundo sabe interpretá-los!

Assim, o cérebro não fica tentando entender o gráfico

É importante que tenham uma linha de **base zero**. Veja a figura para um exemplo ruim



Aumento visual aparente:

- $35 - 34 = 1$ e $39,6 - 34 = 5,6$
- $(5,6 - 1) / 1 = 460\%$

Aumento real:

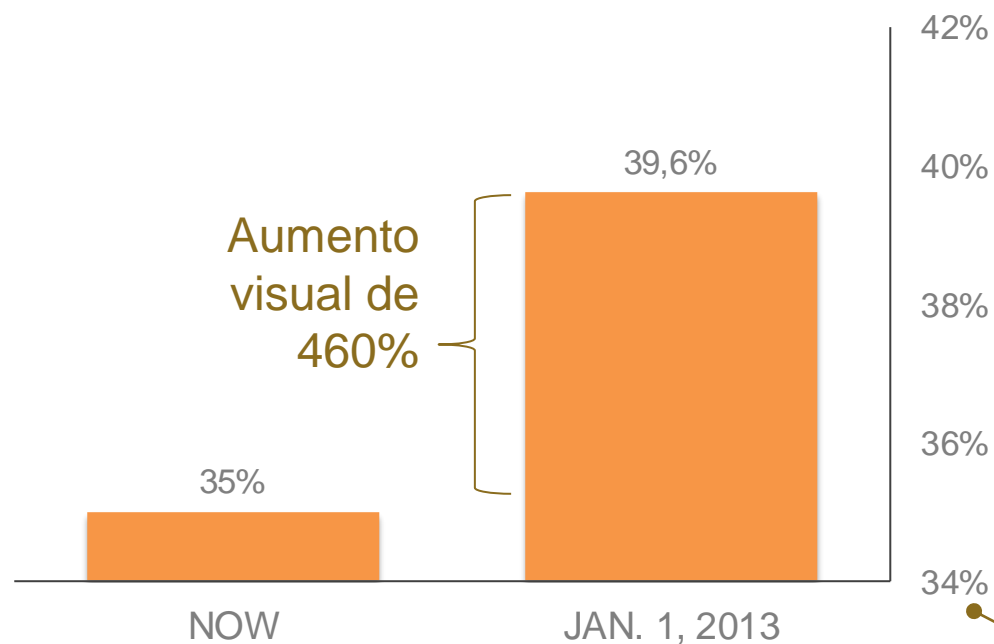
- $35 - 0 = 35$ e $39,6 - 0 = 39,6$
- $(39,6 - 35) / 35 = \sim 0,13 = \sim 13\%$

FIG0213

Non-zero baseline: as originally graphed

IF BUSH TAX CUTS EXPIRE

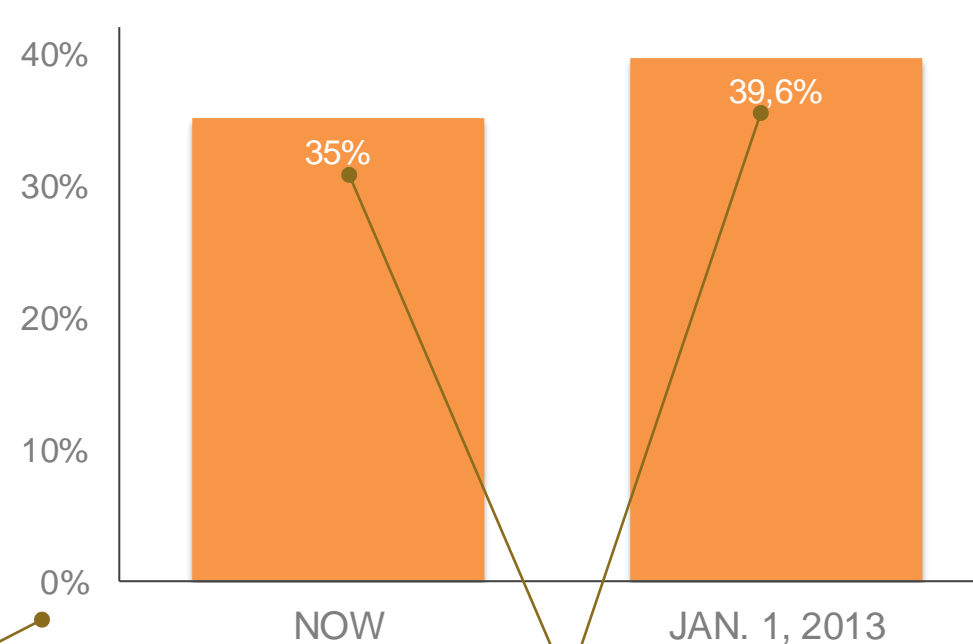
TOP TAX RATE



Zero baseline: as it should be graphed

IF BUSH TAX CUTS EXPIRE

TOP TAX RATE



Legenda y alterada: entender dados antes de ler.
Também poderia omitir

Saturação reduzida

Barras

Dicas:

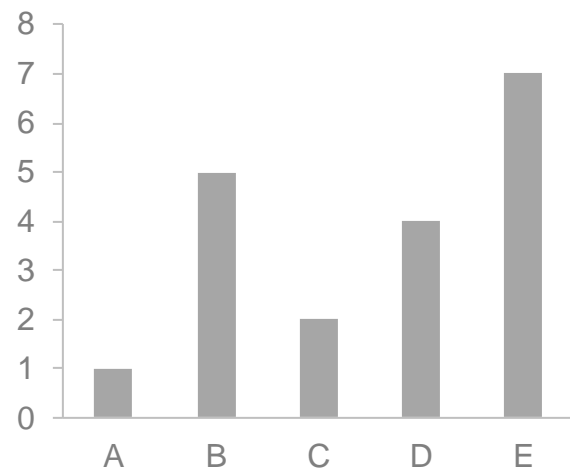
Sempre devem ter uma linha de **base zero**

Cuidado para não distorcer a escala para mostrar o que quer (ética)

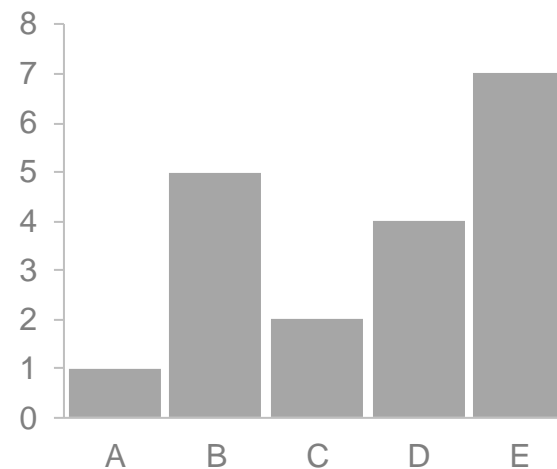
A largura da barra não pode levar a comparação de área (verificar o espaço em branco entre elas)

60214

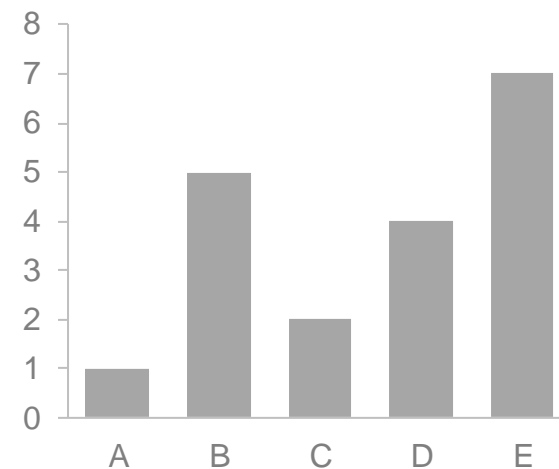
Too thin



Too thick



Just right



Barras

Gráfico de barras verticais

Também conhecido como gráfico de colunas

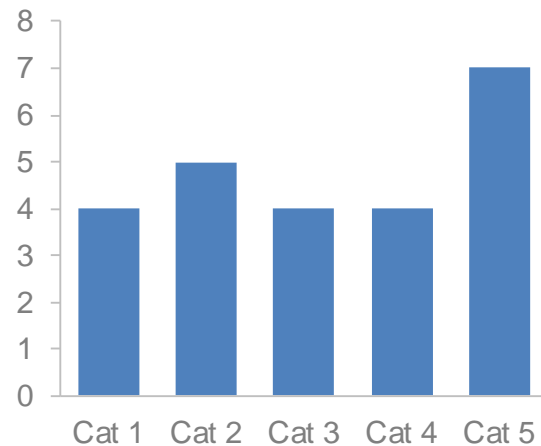
Podem ter várias séries

Quanto mais séries, mais difícil enfocar uma série por vez

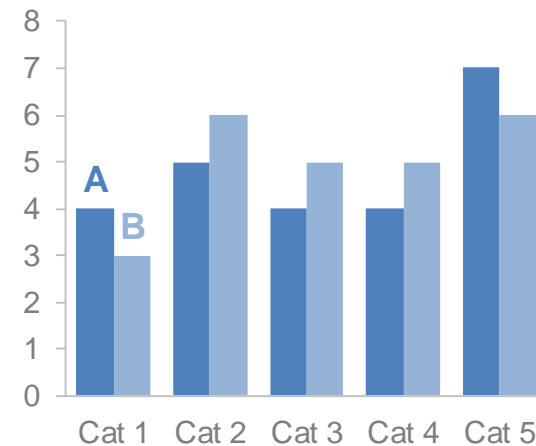
Indicado para comparações com foco na **mesma categoria**

FIG0215

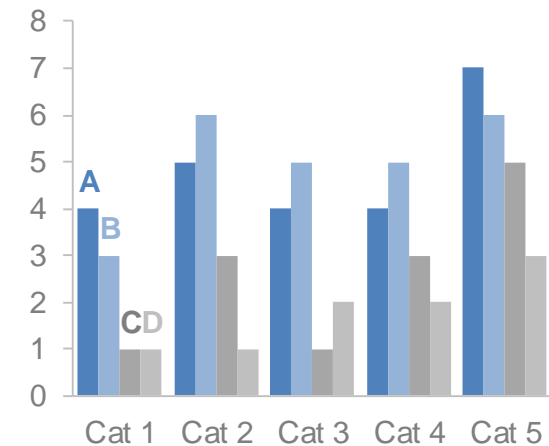
Single series



Two series



Multiple series



Barras

Gráfico de barras verticais empilhadas

Uso mais limitado

Permite comparações de totais entre categorias

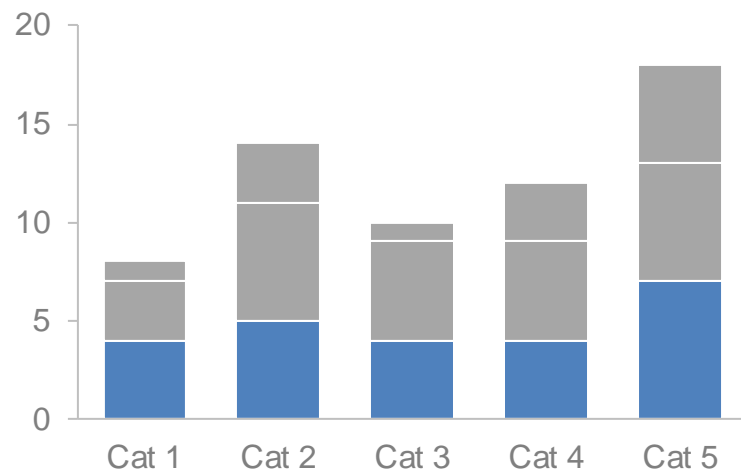
Limitante: Variação de cores nos aplicativos de geração

Mais difícil comparar após ultrapassar a série inferior

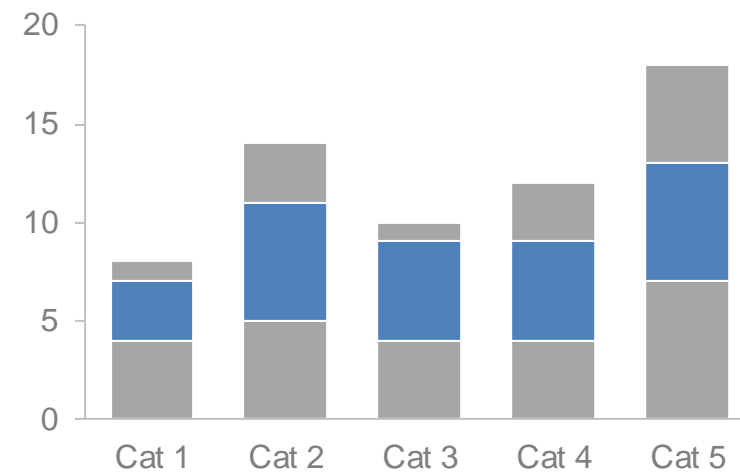
Há uma variação com barras empilhadas a 100%

FIG0216

Comparing **these** is easy



Comparing **these** is hard



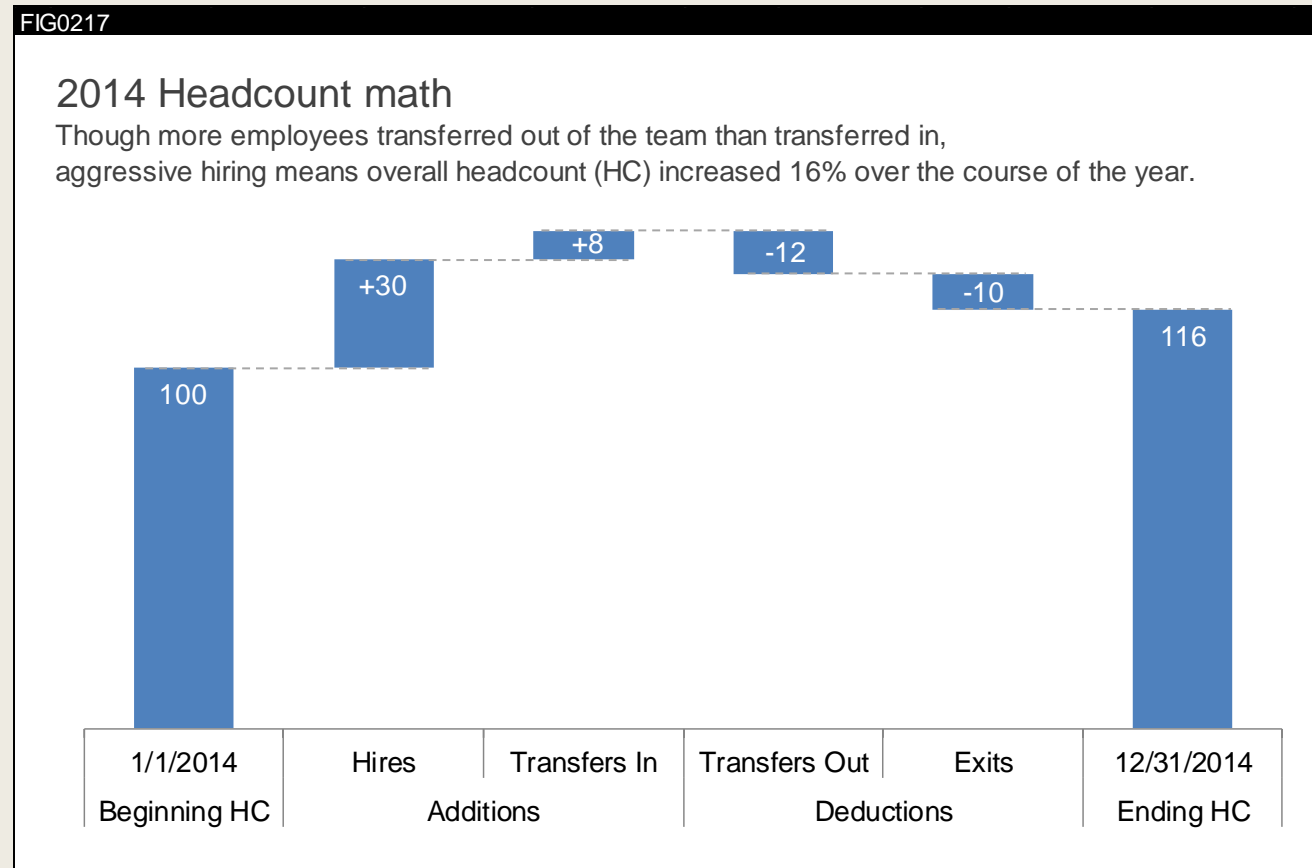
Barras

Gráfico de cascata

Decomposição das partes de um gráfico barras empilhadas

Não faz parte dos gráficos padrão do Excel. Guia:

<https://www.storytellingwithdata.com/blog/2011/11/waterfall-chart>



Barras

Gráfico de barras horizontais

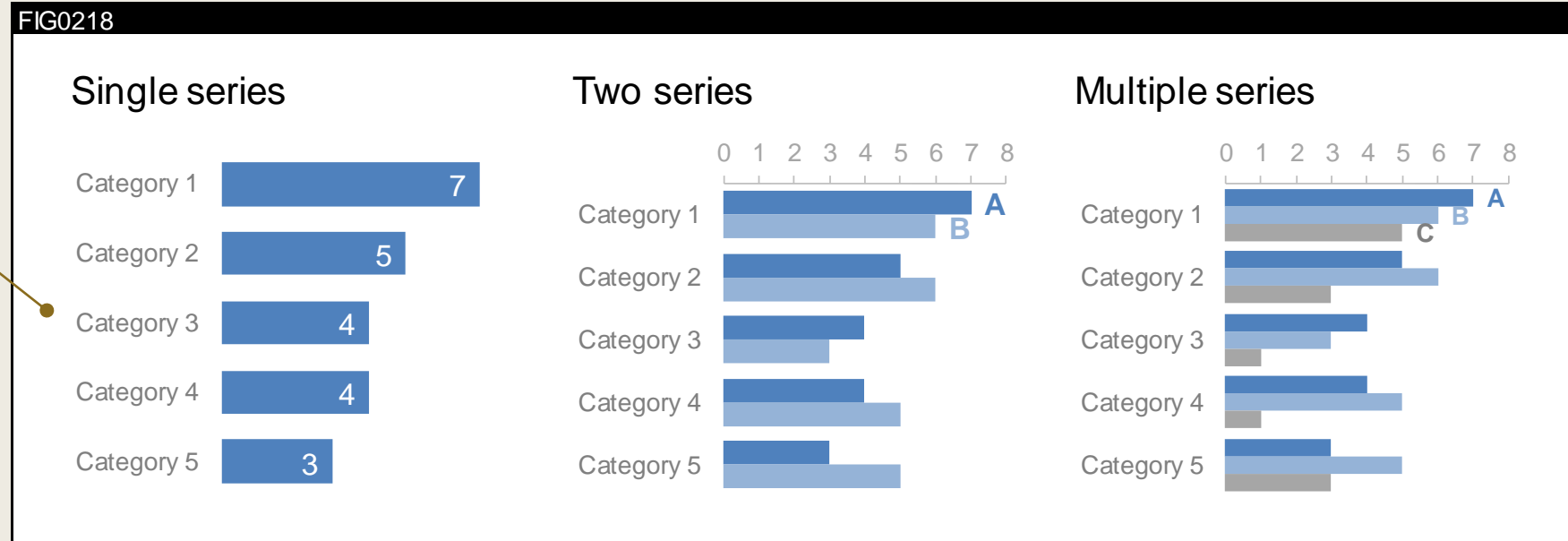
Excelente opção para **dados categóricos**

Permite **conteúdo maior** na legenda das categorias

Segue o fluxo em Z de leitura: **nome das categorias antes dos dados**

Ordenação das categorias.
Lembre-se do fluxo Z:

- Ordem natural (p. ex. faixa etária);
- Ou, o mais importante primeiro



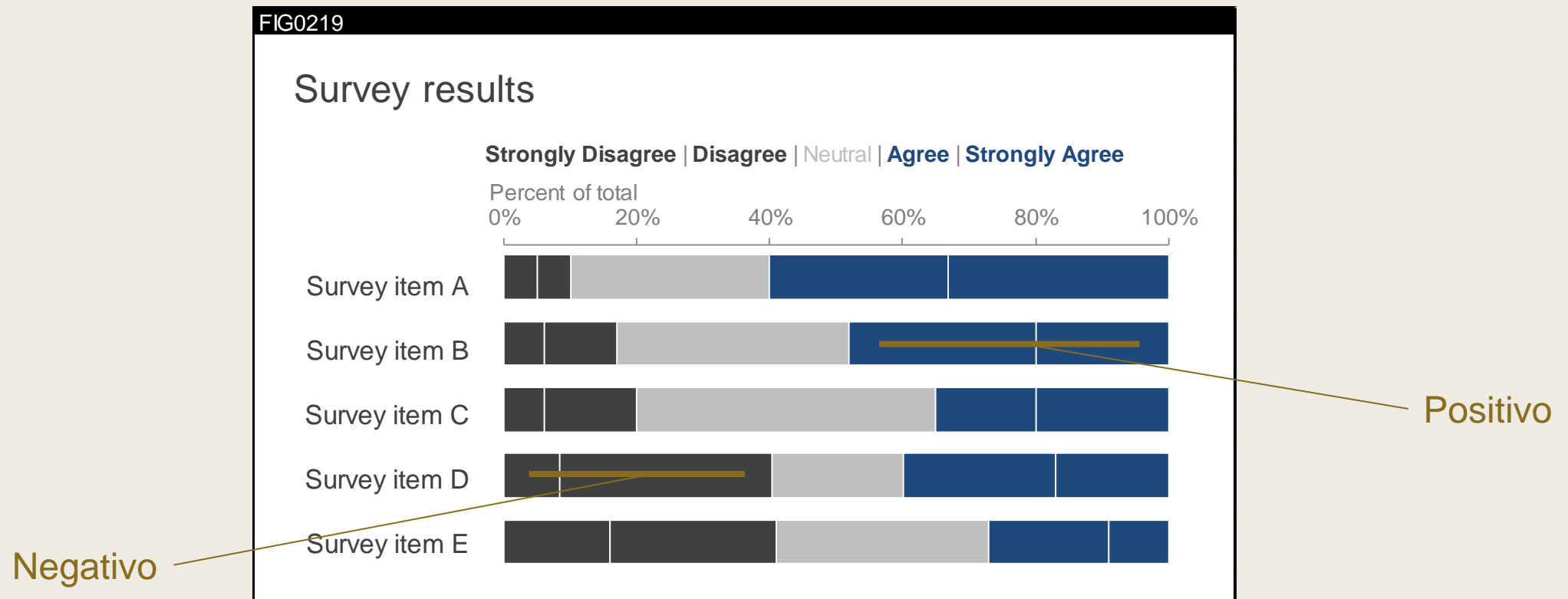
Barras

Gráfico de barras horizontais empilhadas

Pode mostrar valores **absolutos** ou a **soma até 100%**

Visualização adequada para escala do **negativo** para o **positivo**
(agrupada)

Veja agrupamento abaixo de uma escala **Likert**

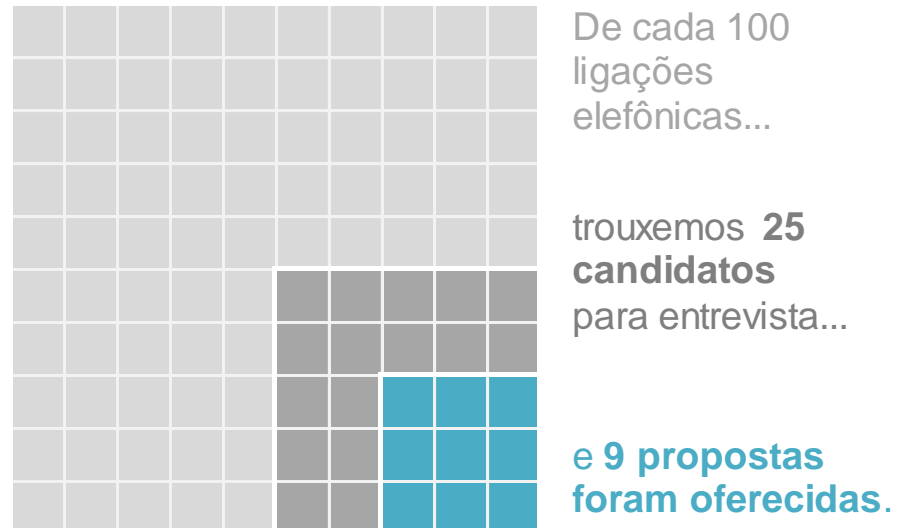


Área

Olhos humanos não atribuem bem valor quantitativo no espaço bidimensional

FIG0220

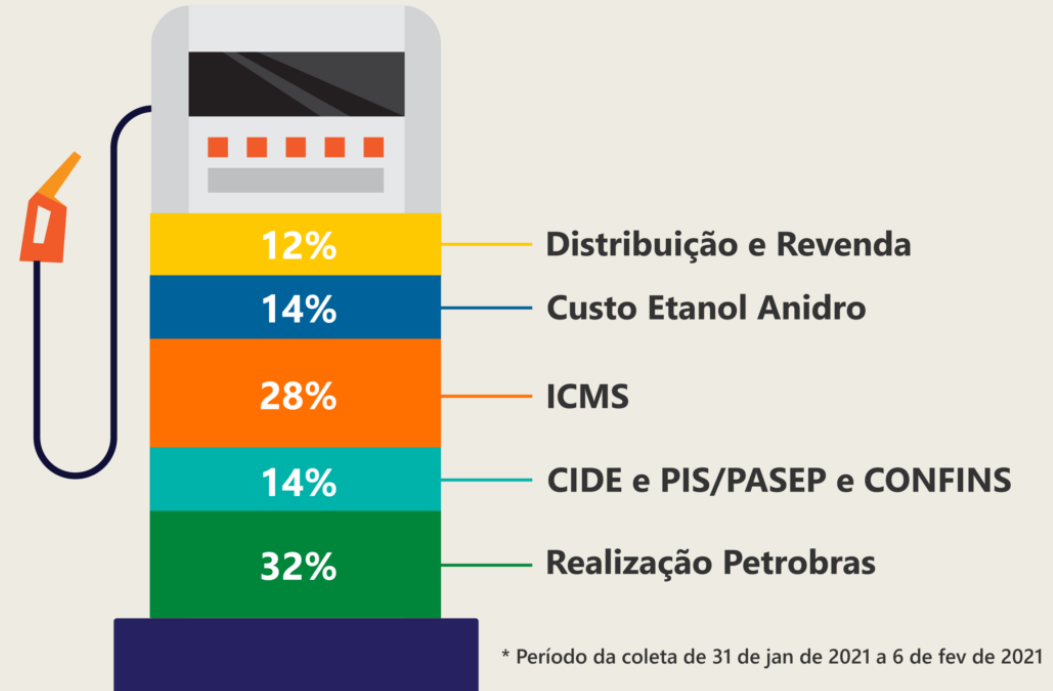
Decomposição da entrevista




Infográficos

Representação gráfica de informações ou dados

Deve-se evitar distorcer, super ou subdimensionar os dados





Gráficos a serem evitados

A ser evitado

Gráficos de pizza

Gráficos de rosca

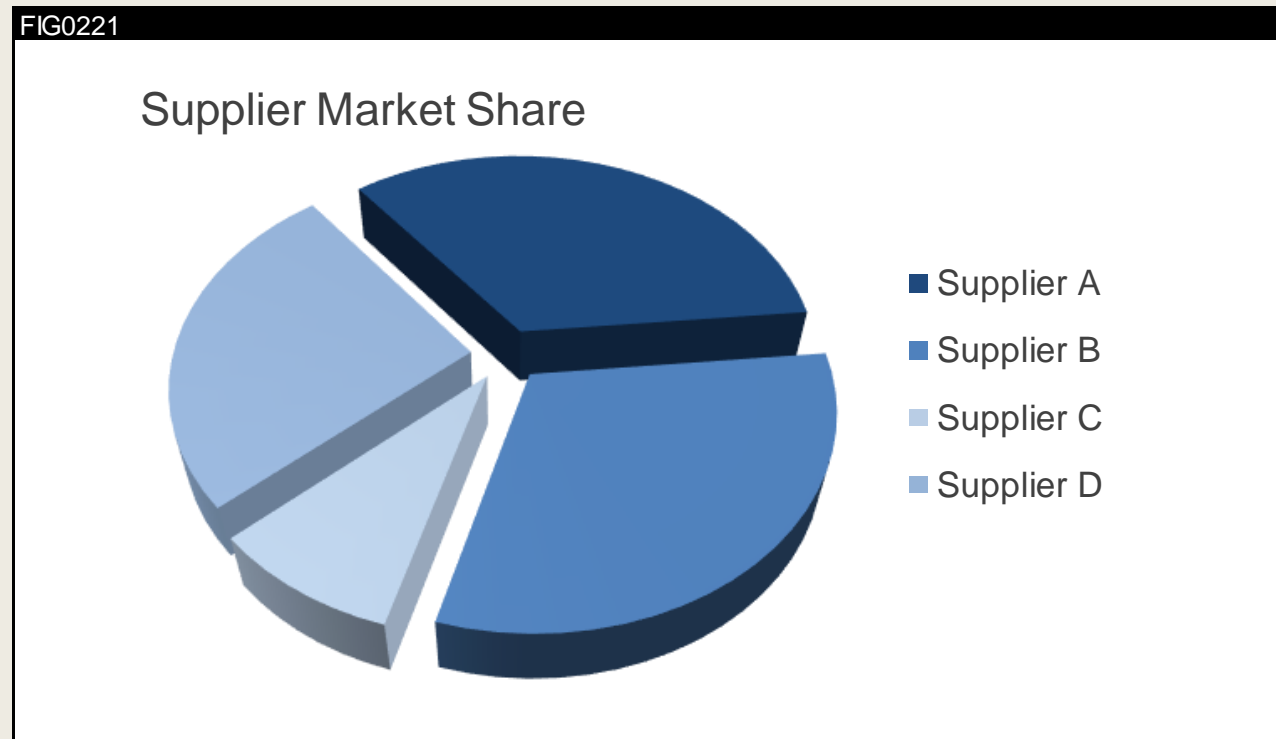
Gráficos 3D

Gráficos com eixo y secundário



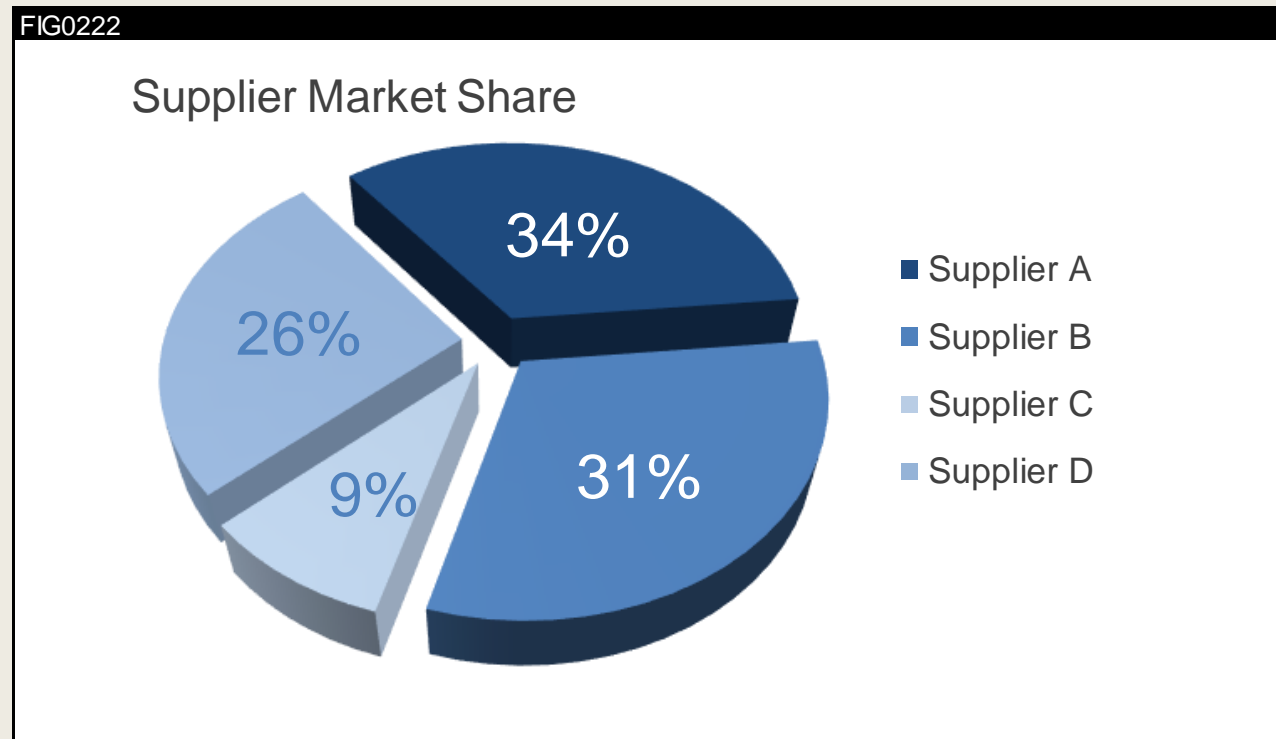
Gráficos de pizza

Qual seria o maior fornecedor? Fornecedor A ou B?



Gráficos de pizza

A perspectiva dos gráficos 3D **distorce** a percepção visual
Mesmo os bidimensionais (pizza tradicional) oferecem desafios!



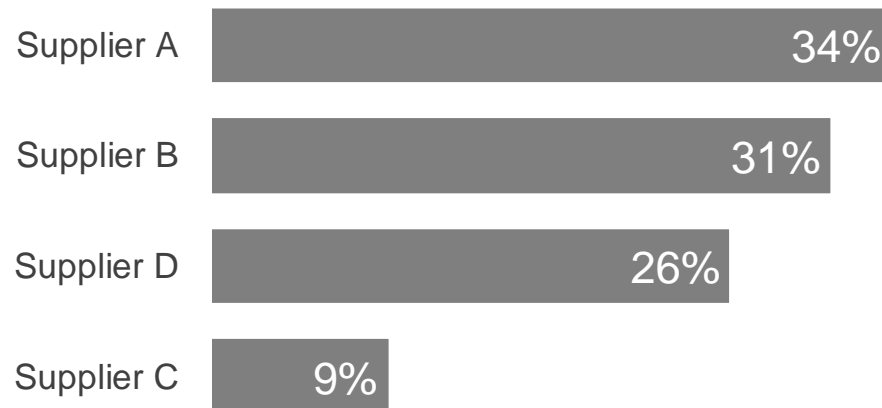
Assim fica melhor!

Aproveita-se a fluxo de leitura em Z

Linha base comum: fácil avaliar o tamanho relativo

FIG0223

Supplier Market Share

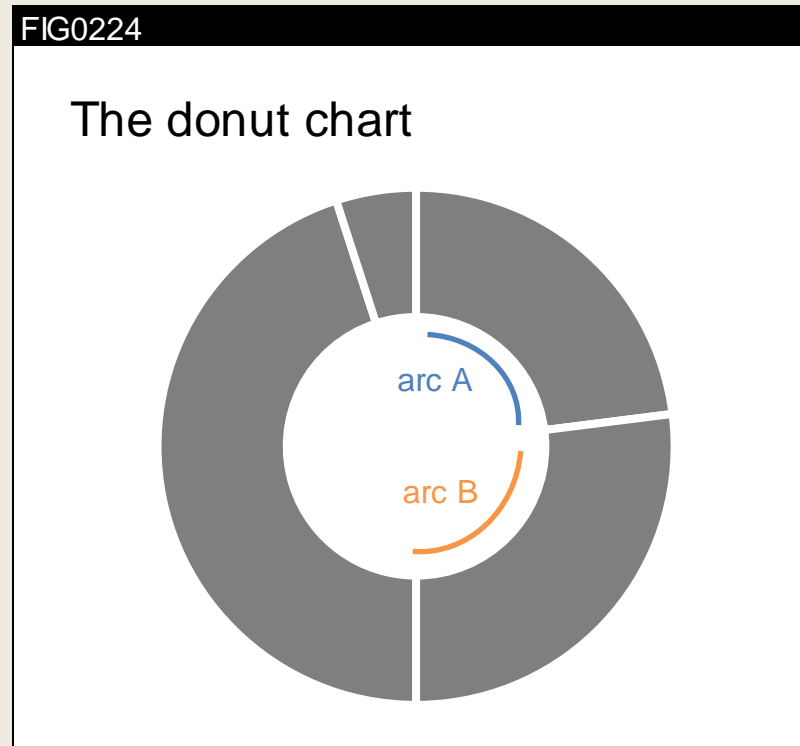


Tenta
representar o
efeito da pizza
em mostrar o
todo 100%

Gráficos de rosca

Na pizza pedimos ao público para comparar **ângulo** e **área**

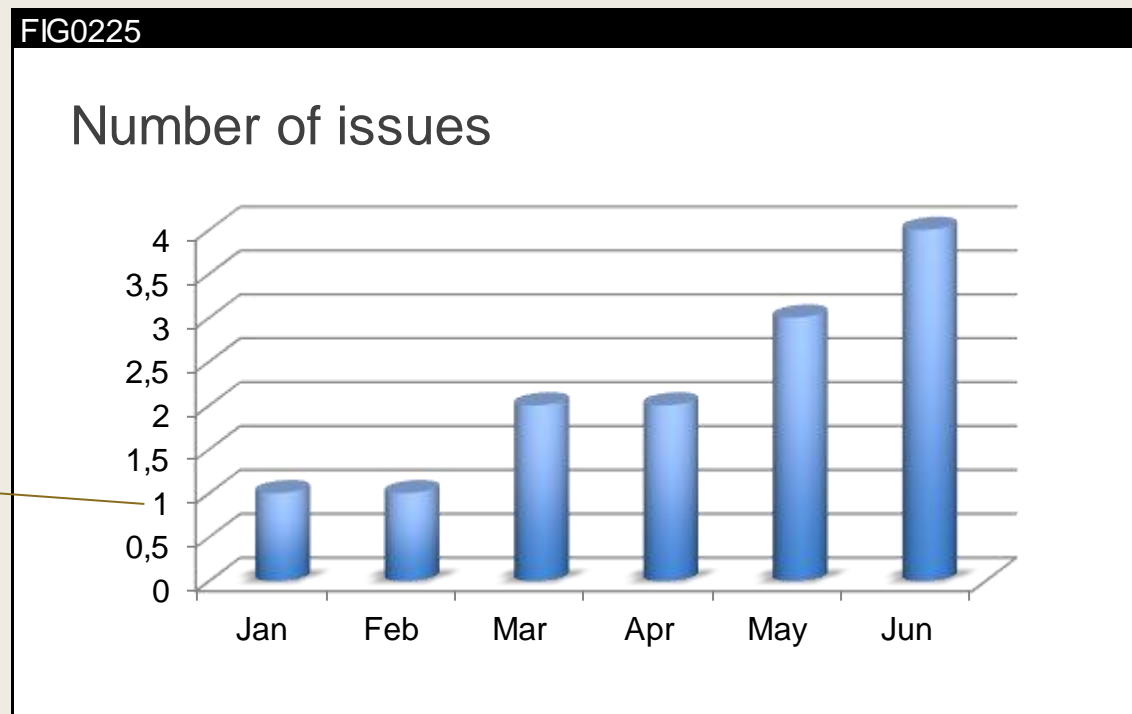
Aqui pedimos ao público para comparar o **comprimento** de um arco!



Mais um exemplar em 3D

Quantas publicações existem em janeiro e fevereiro?

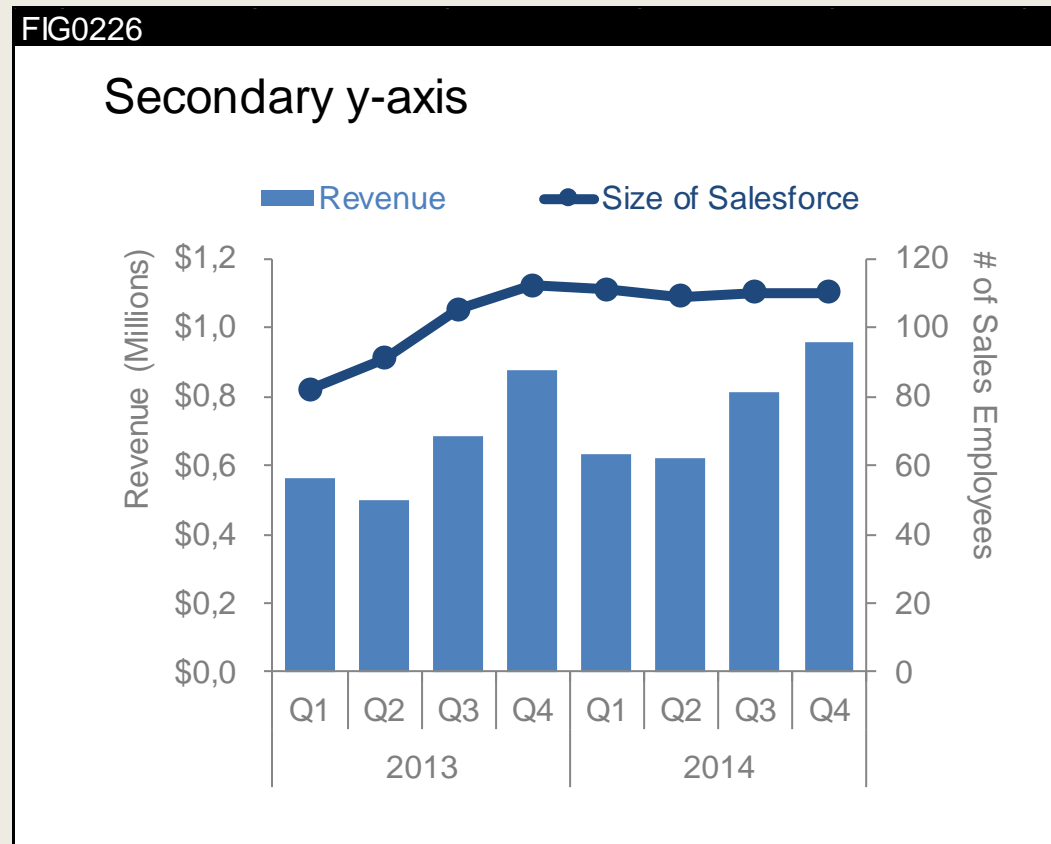
Nos dados do gráfico temos 1 publicação para ambos os meses!



Gráficos com eixo y secundário

Demanda tempo e leitura para entender quais dados sob quais eixos

Deve-se evitar este tipo de gráfico

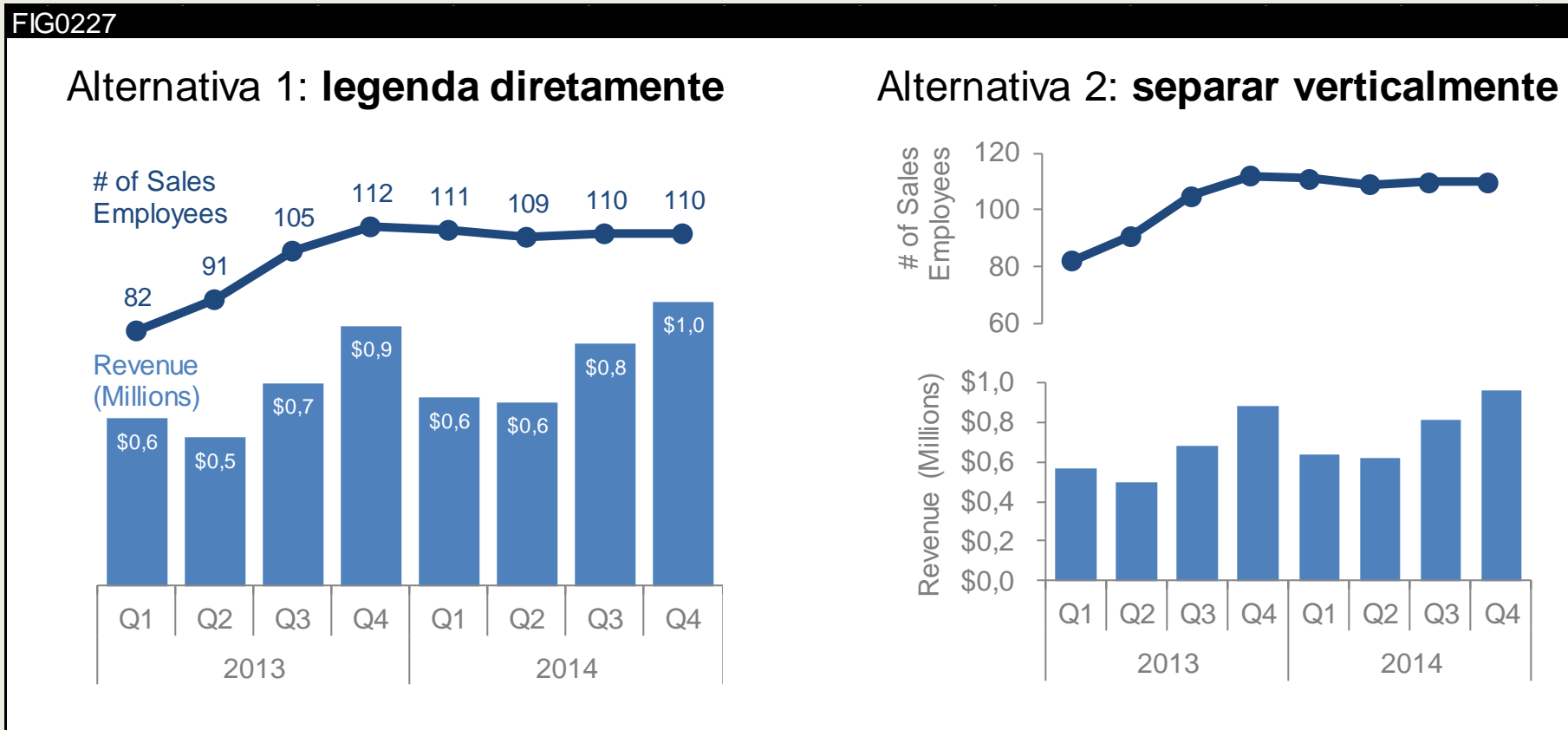


Gráficos com eixo y secundário

Alternativas

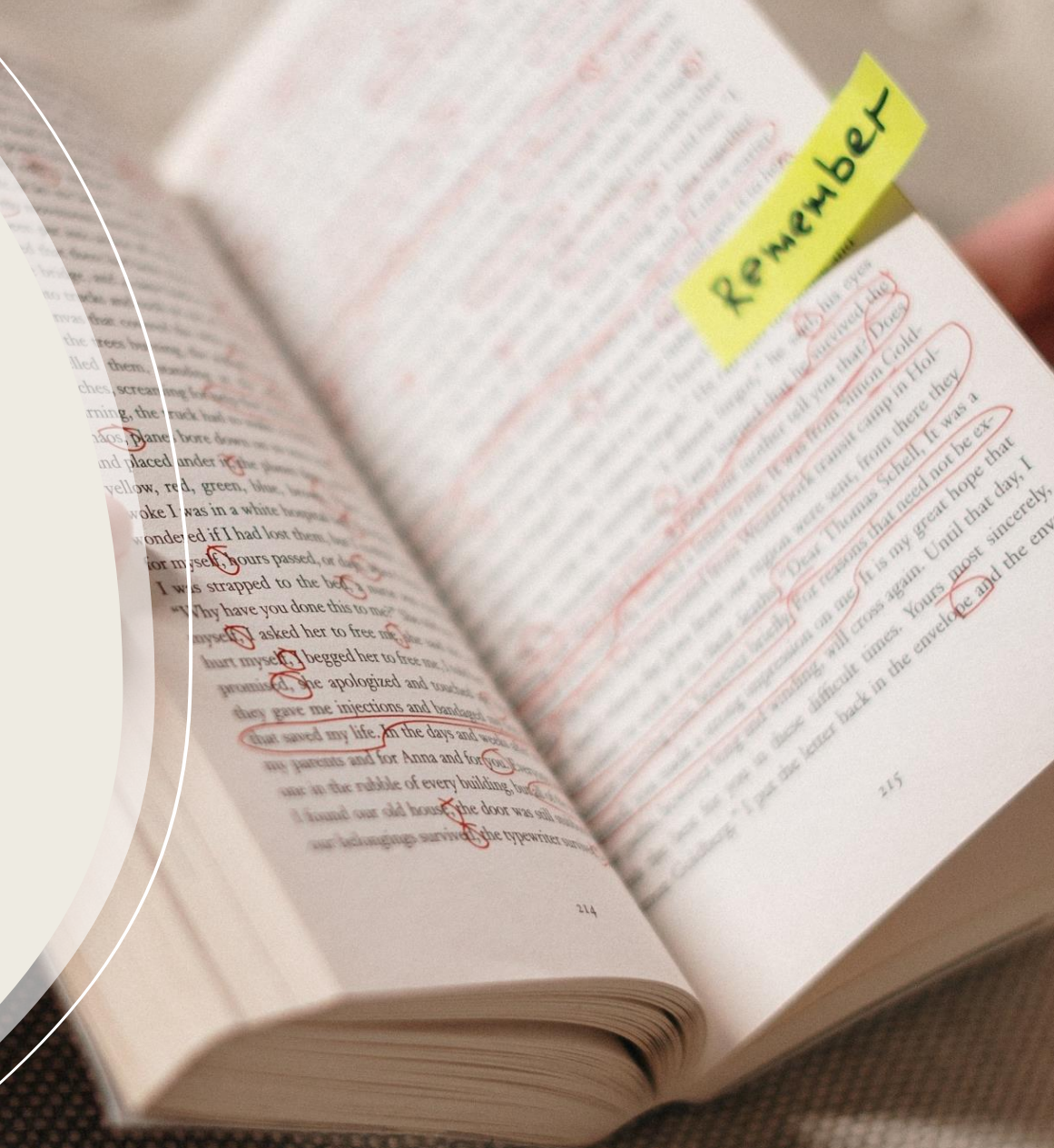
- Focos:
- Alternativa 1: atenção aos números
 - Alternativa 2: focaliza a tendência dominante

Dois conjuntos de dados no mesmo eixo (x): verificar se a **relação** existe



Para não esquecer

- **Poucos tipos** de representações visuais para a **maioria das necessidades**
- A escolha vai depender daquilo que **você precisa** que seu **público saiba**
- Na incerteza de escolher o **gráfico adequado**, mostre a um colega e peça **feedback**





Vamos praticar

Melhore esta tabela

PASSO 1. Observações sobre a tabela? Há alguma dúvida sobre os dados?

PASSO 2. Vamos assumir que a comunicação acontecerá por meio de uma tabela. Como seria sua versão aprimorada?

PASSO 3. Vamos mostrar como as contas estão distribuídas comparando-as como as receitas. Crie um gráfico.

| New client tier share | | | | |
|-----------------------|---------------|------------|---------------|-----------|
| Tier | # of Accounts | % Accounts | Revenue (\$M) | % Revenue |
| A | 77 | 7,08% | \$4,68 | 25% |
| A+ | 19 | 1,75% | \$3,93 | 21% |
| B | 338 | 31,07% | \$5,98 | 32% |
| C | 425 | 39,06% | \$2,81 | 15% |
| D | 24 | 2,21% | \$0,37 | 2% |

PASSO 1. Observações sobre a tabela? Há alguma dúvida sobre os dados?

Precisão de 2 casas é realmente necessária?
Faz diferença 7,08 de 7,09?

New client tier share

| Tier | # of Accounts | % Accounts | Revenue (\$M) | % Revenue |
|------|---------------|------------|---------------|-----------|
| A | 77 | 7,08% | \$4,68 | 25% |
| A+ | 19 | 1,75% | \$3,93 | 21% |
| B | 338 | 31,07% | \$5,98 | 32% |
| C | 425 | 39,06% | \$2,81 | 15% |
| D | 24 | 2,21% | \$0,37 | 2% |

Maioria das
contas

Maior receita

Segmentos
estão em
ordem?

Falta uma linha
de total. Chegam
a 100%?

PASSO 2. Vamos assumir que a comunicação acontecerá por meio de uma tabela. Como seria sua versão aprimorada?

VERSÃO COM LINHAS SOMBREADAS

Ordenação
para fazer
sentido

Sombras
nas linhas

| New client tier share | | | | |
|-----------------------|---------------|------------|---------------|-----------|
| Tier | # of Accounts | % Accounts | Revenue (\$M) | % Revenue |
| A+ | 19 | 2% | \$3,9 | 21% |
| A | 77 | 7% | \$4,7 | 25% |
| B | 338 | 31% | \$6,0 | 32% |
| C | 425 | 39% | \$2,8 | 15% |
| D | 24 | 2% | \$0,4 | 2% |
| All other | 205 | 19% | \$0,9 | 5% |
| TOTAL | 1.088 | 100% | \$18,7 | 100% |

Total a 100%

VERSÃO COM MAPA DE CALOR

Colunas
mescladas

| New client tier share | | | | |
|-----------------------|----------|----------|---------|----------|
| TIER | ACCOUNTS | | REVENUE | |
| | # | % OF TOT | \$M | % OF TOT |
| A+ | 19 | 2% | \$3,9 | 21% |
| A | 77 | 7% | \$4,7 | 25% |
| B | 338 | 31% | \$6,0 | 32% |
| C | 425 | 39% | \$2,8 | 15% |
| D | 24 | 2% | \$0,4 | 2% |
| All other | 205 | 19% | \$0,9 | 5% |
| TOTAL | 1.088 | 100% | \$18,7 | 100% |

PASSO 2. Vamos assumir que a comunicação acontecerá por meio de uma tabela. Como seria sua versão aprimorada?

Fluxo Z: o mais importante no topo à esquerda

VERSÃO COM BARRAS EMBUTIDAS

| New client tier share | | | | |
|-----------------------|----------|-------------|---------|-------------|
| TIER | ACCOUNTS | | REVENUE | |
| | # | % OF TOT | \$M | % OF TOT |
| A+ | 19 | <div></div> | \$3,9 | <div></div> |
| A | 77 | <div></div> | \$4,7 | <div></div> |
| B | 338 | <div></div> | \$6,0 | <div></div> |
| C | 425 | <div></div> | \$2,8 | <div></div> |
| D | 24 | <div></div> | \$0,4 | <div></div> |
| All other | 205 | <div></div> | \$0,9 | <div></div> |
| TOTAL | 1.088 | 100% | \$18,7 | 100% |

Poderia inverter as colunas, caso receita fosse mais importante.
Fluxo Z!

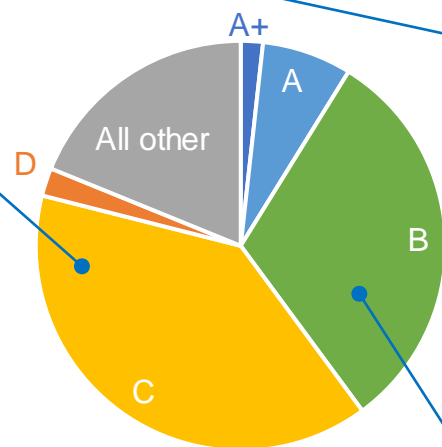
PASSO 3. Vamos mostrar como as contas estão distribuídas comparando-as como as receitas. Crie um gráfico.

Segmentos estão em diferentes locais
(dificulta a comparação)

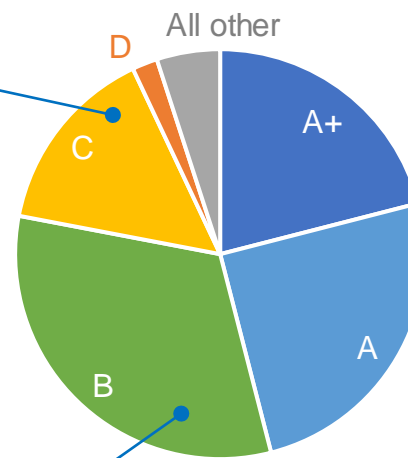
VERSÃO PIZZA “PARTE DE UM TODO”

New client tier share

% of Total **Accounts**



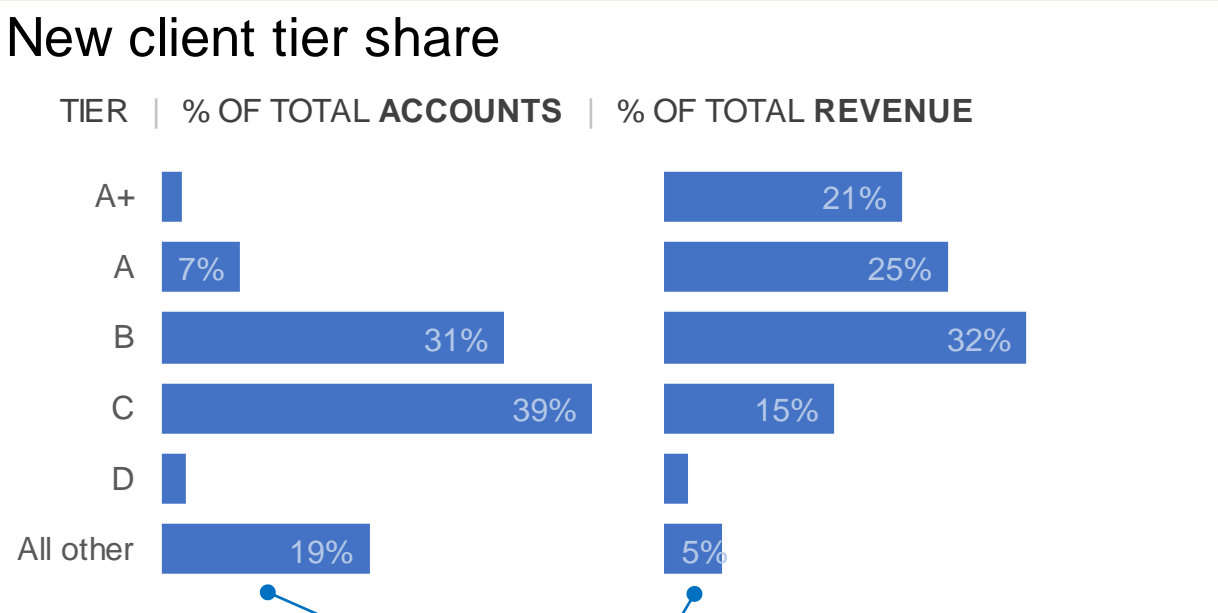
% of Total **Revenue**



Difícil distinguir as áreas
(e ângulos)

PASSO 3. Vamos mostrar como as contas estão distribuídas comparando-as como as receitas. Crie um gráfico.

VERSÃO COM DOIS GRÁFICOS DE BARRAS HORIZONTAIS

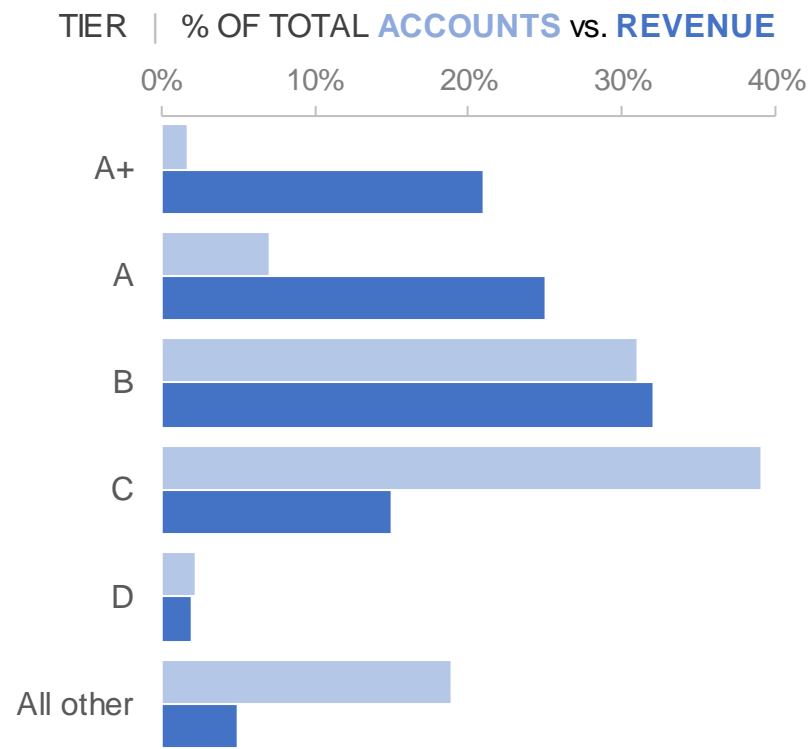


A comparação ainda
poderá ser melhorada

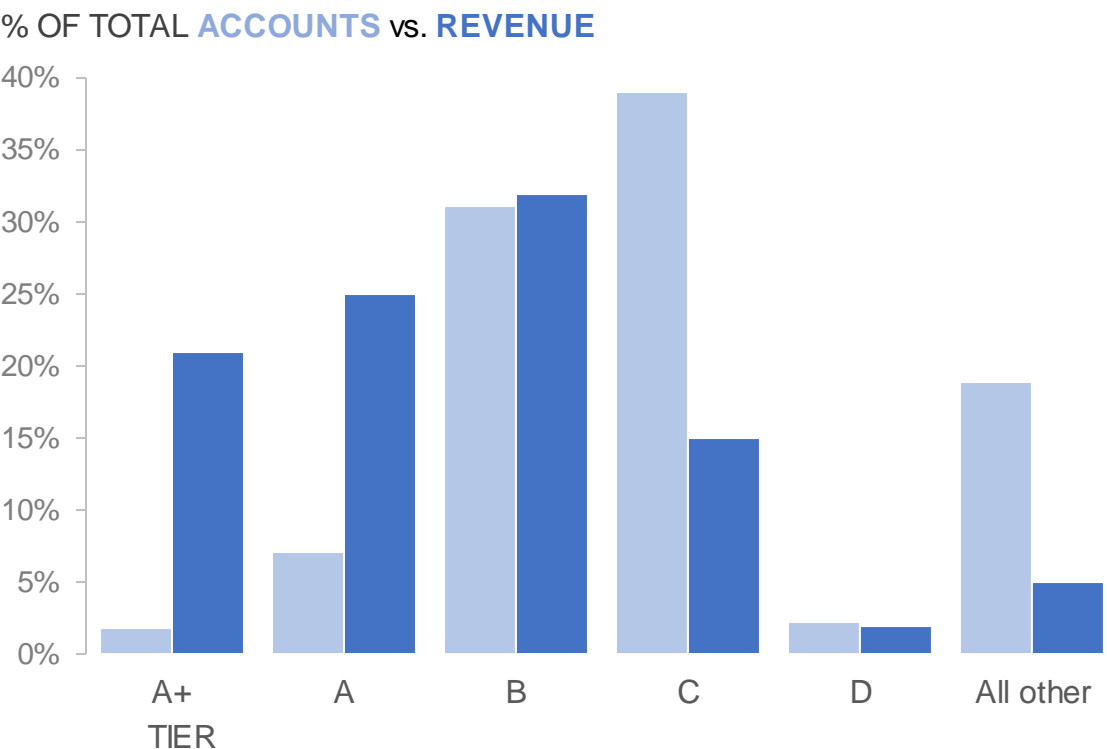
PASSO 3. Vamos mostrar como as contas estão distribuídas comparando-as como as receitas. Crie um gráfico.

VERSÕES PARA COMPARAÇÃO NUMA MESMA BASE

New client tier share

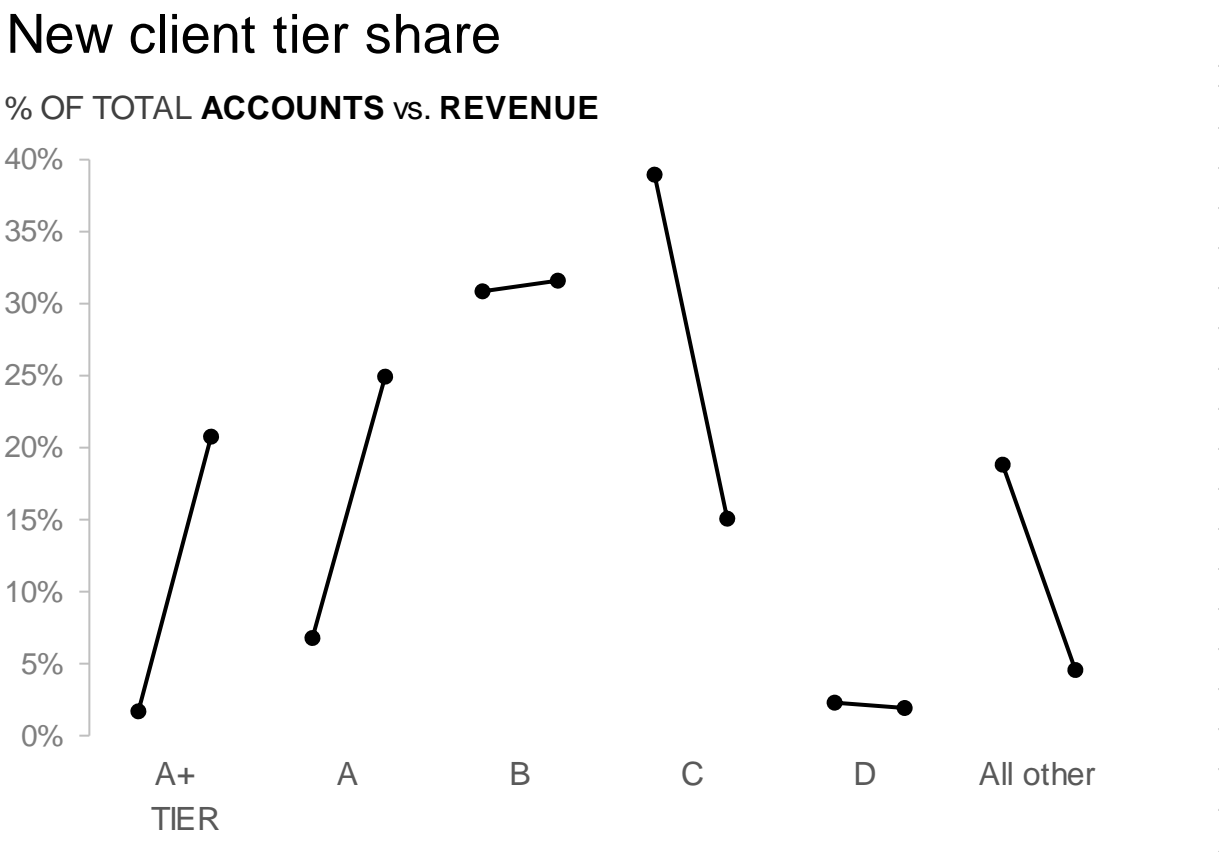
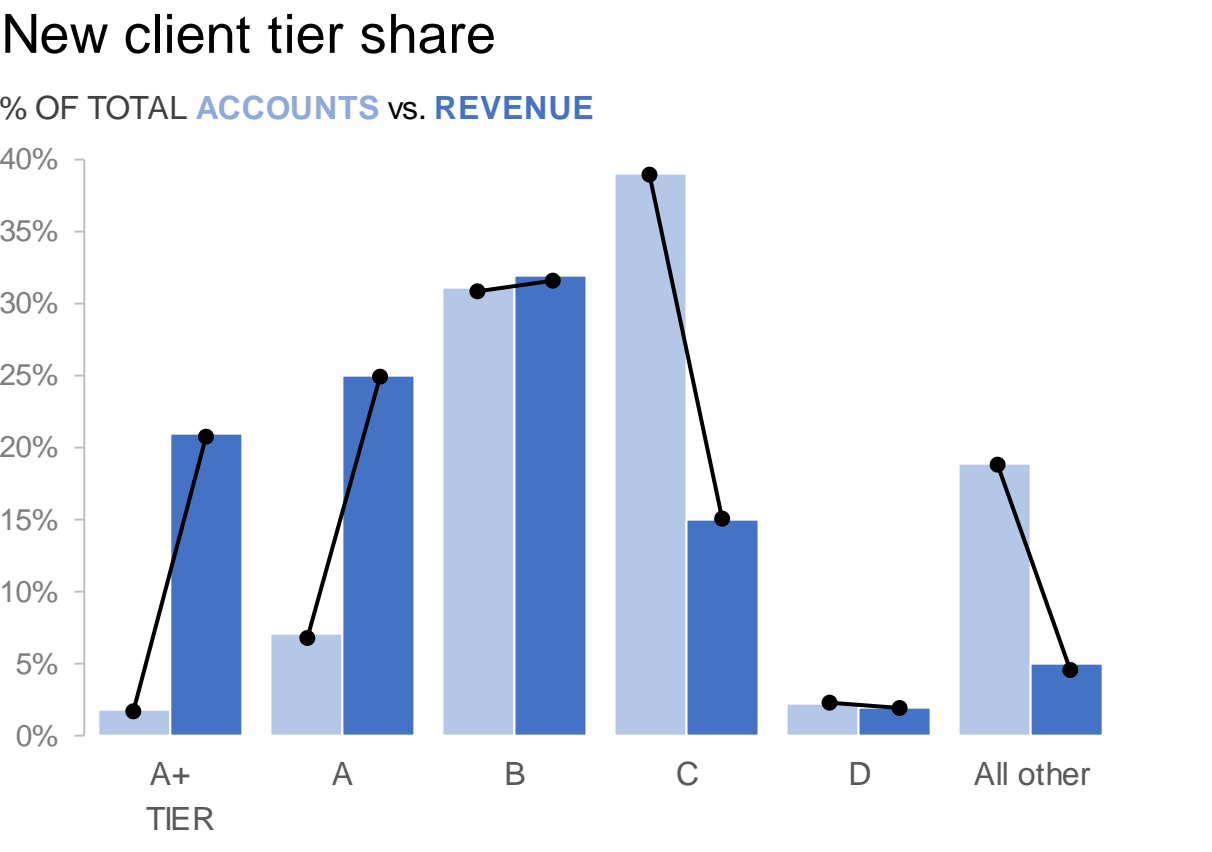


New client tier share



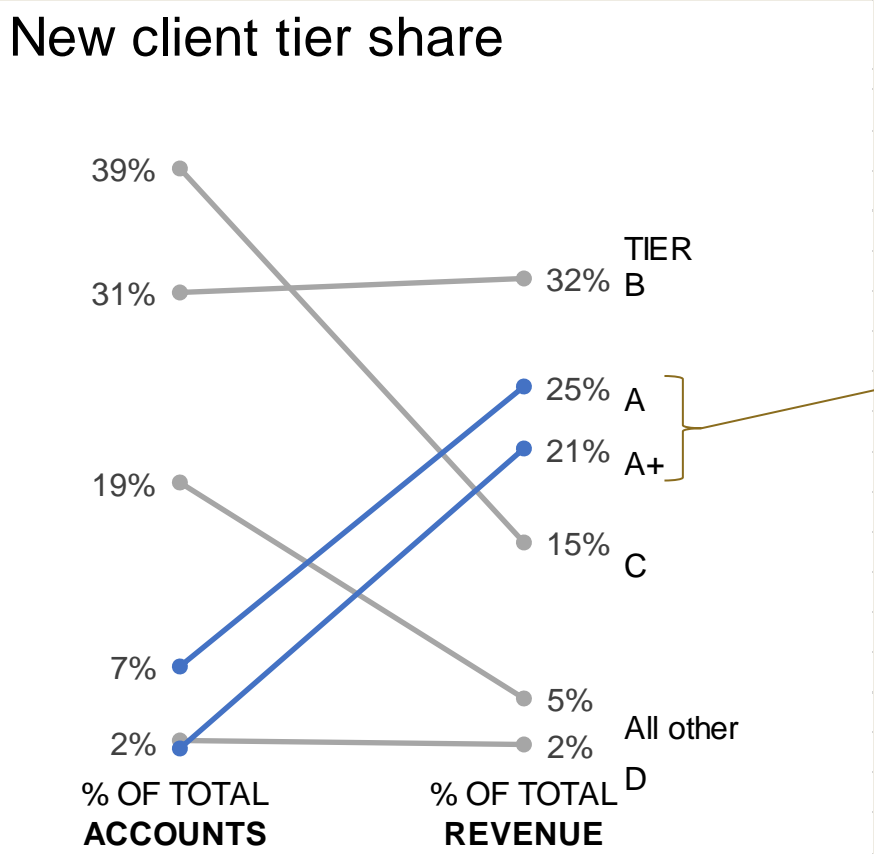
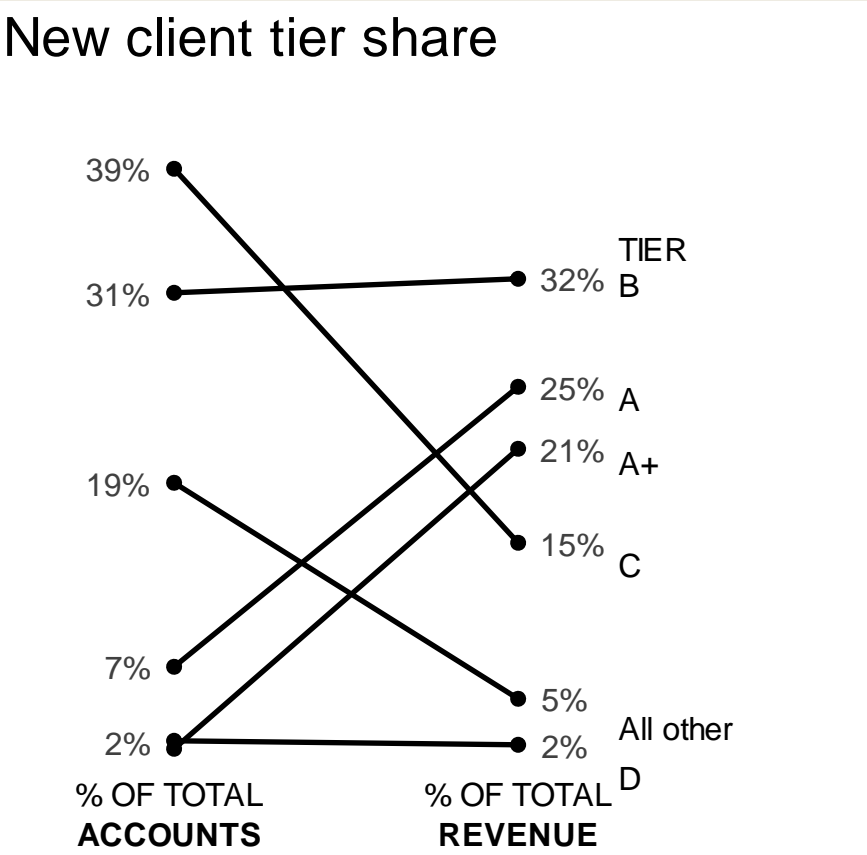
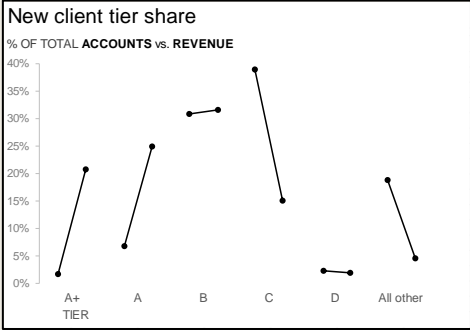
PASSO 3. Vamos mostrar como as contas estão distribuídas comparando-as como as receitas. Crie um gráfico.

VERSÃO COM O DEPICT DAS COMPARAÇÕES
(MESMO PROCESSO QUE NOSSOS OLHOS FAZEM INTUITIVAMENTE)



PASSO 3. Vamos mostrar como as contas estão distribuídas comparando-as como as receitas. Crie um gráfico.

VERSÃO COM UM GRÁFICO DE INCLINAÇÃO



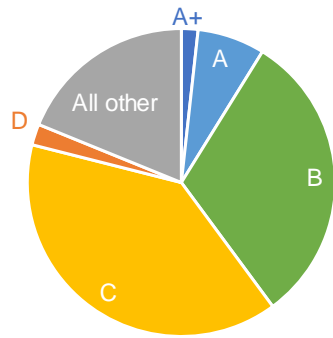
Enfatizando a participação das camadas A+ e A na receita:

- 9% das contas
- 46% de receita

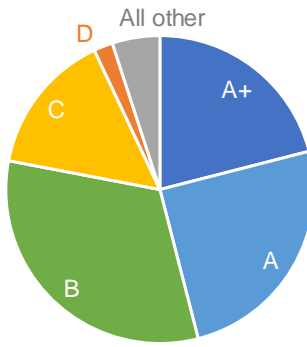
Qual gráfico você mais gostou?

New client tier share

% of Total **Accounts**

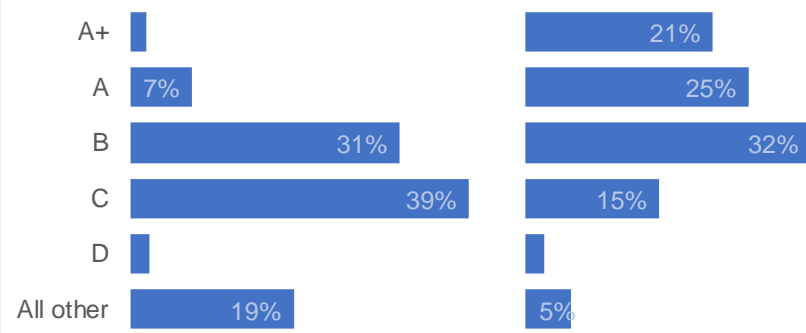


% of Total **Revenue**



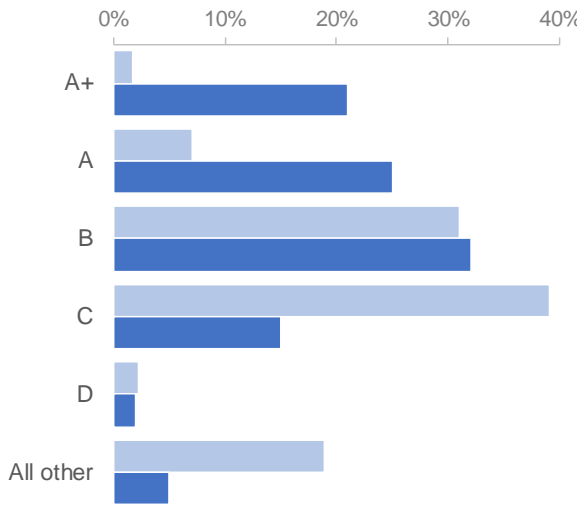
New client tier share

TIER | % OF TOTAL **ACCOUNTS** | % OF TOTAL **REVENUE**



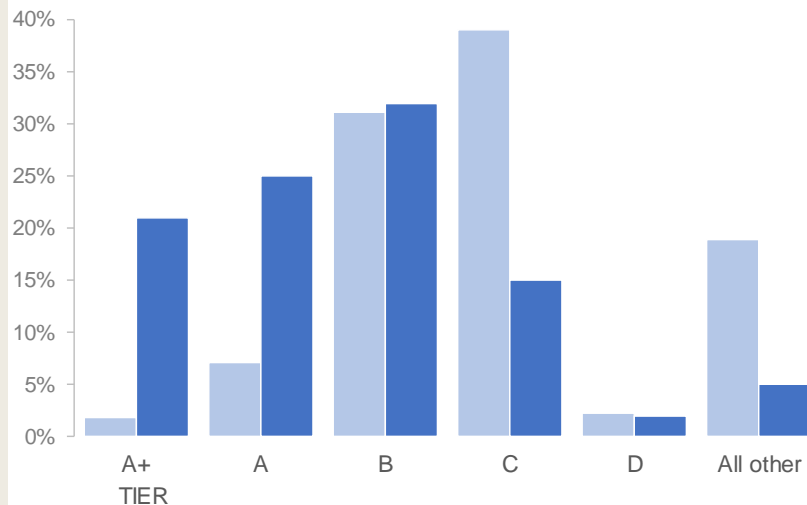
New client tier share

TIER | % OF TOTAL **ACCOUNTS** vs. **REVENUE**

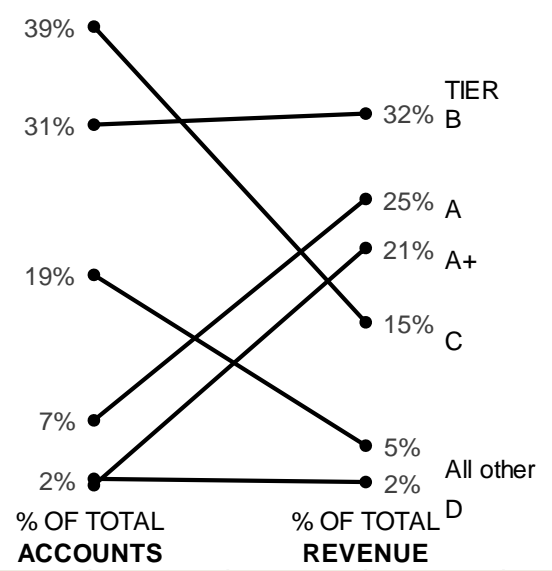


New client tier share

% OF TOTAL **ACCOUNTS** vs. **REVENUE**



New client tier share



Créditos

Bibliografia:

KNAFLIC, Cole N. Storytelling com dados: um guia sobre visualização de dados para profissionais de negócios. Rio de Janeiro: Alta Books, 2017.

KNAFLIC, Cole N. Storytelling com dados: Let's practice. Wiley, 2019.

Fotos:

<https://www.pexels.com/pt-br/>

Cliparts:

<https://www.gratispng.com/>