



Modelagem, Extração e Manipulação de Dados

BLOCO: B.I. E ANÁLISE DE DADOS

PROF. RODRIGO EIRAS, M.SC.

[ETAPA 4] AULAS 1 E 2 – DASHBOARDS E VISUALIZAÇÕES DE DADOS



Na última aula...

- Power BI
- Tipos de Fontes de Dados
- Exemplos de Fontes de Dados
- Conexões Remotas com MySQL Server

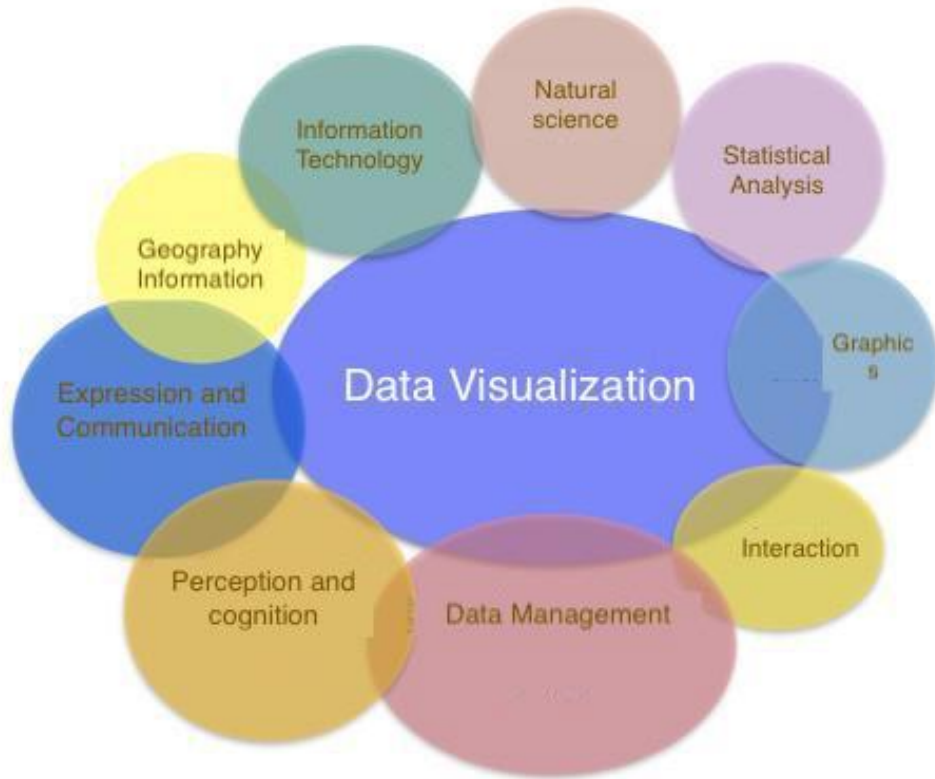


Agenda

- Visualização de Dados
- Criar relatórios no Power BI
- Classificações e Ordenações
- Arredondamentos

Porque visualização de dados é importante?





Porque visualização de dados é importante?

- As ferramentas de Visualização de dados provêm uma visualização interativa vai um passo além da exibição de gráficos estáticos para fazer uso de funcionalidades (*drill-down* e *drill-up*) que dão interatividade e dinamismo na análise de dados.
- A visualização de dados do contexto de Big Data trata da apresentação dos dados em um formato gráfico e mais lúdico.
- Vamos começar apresentando como criar relatórios no Power BI.
- É muito fácil e intuitivo, o Power BI é destaque no mercado também na categoria usabilidade.

Tabelas

	A	B	C	D
1	Produto ▼	Trim 1 ▼	Trim 2 ▼	Total Gera ▼
2	Chocolate	R\$ 744,60	R\$ 162,56	R\$ 907,16
3	Bala de goma	R\$ 5.079,60	R\$ 1.249,20	R\$ 6.328,80
4	Pão escocês	R\$ 1.267,50	R\$ 1.062,50	R\$ 2.330,00
5	Scones do Sr. Diogo	R\$ 1.418,00	R\$ 756,00	R\$ 2.174,00
6	Torta de açúcar	R\$ 4.728,00	R\$ 4.547,92	R\$ 9.275,92
7	Biscoitos de chocolate	R\$ 943,89	R\$ 349,60	R\$ 1.293,49
8	Total	R\$ 14.181,59	R\$ 8.127,78	R\$ 22.309,37

- As tabelas são importantes formas de representação de dados estatísticos, pelo seu aspecto quantitativo e maior facilidade na investigação pretendida.
- Podemos utilizar as tabelas para:
 - organizar dados
 - visualizar melhor os resultados obtidos numa pesquisa
 - comparar quantidades
 - elaborar estatísticas

TABELA 6

Distribuição numérica das fêmeas de *Culex (Melanoconion)*, obtidas nos vários métodos de coleta (período de fevereiro a novembro de 1986).

Métodos de coleta	Aspiração	Rede Manual	Isca Humana	Total
Espécies				
<i>Cx. aliciae</i>	8	—	2	10
<i>Cx. aureonotatus</i>	12	—	2	14
<i>Cx. bastagarius</i>	52	—	2	54
<i>Cx. delpontei</i>	17	—	1	18
<i>Cx. evansae</i>	33	—	3	36
<i>Cx. lopesi</i>	7	2	1	10
<i>Cx. misionensis</i>	5	1	1	7
<i>Cx. ocellatus</i>	2	2	—	4
<i>Cx. ocosa</i>	11	1	1	13
<i>Cx. plectoporce</i>	99	—	—	99
<i>Cx. (Mel.) sp. grupo atratus</i>	15	2	5	22
<i>Cx. (Mel.) sp. grupo intricatus</i>	54	8	11	73
<i>Cx. (Mel.) sp. grupo pilosus</i>	123	24	1	148
<i>Cx. spissipes</i>	41	12	14	67
<i>Cx. taeniopus</i>	9	6	51	66
<i>Cx. theobaldi</i>	5	1	3	9
<i>Cx. zeteki</i>	41	—	6	47
Outras*	620	70	2.169	2.859
Total	1.154	129	2.273	3.556

* *Cx. oedipus*; *Cx. pedroi*; *Cx. ribeirensis*; *Cx. sacchettae*; *Cx. (Melanoconion) sp.*
Até o mínimo total de 100 exemplares no período.

Tabelas

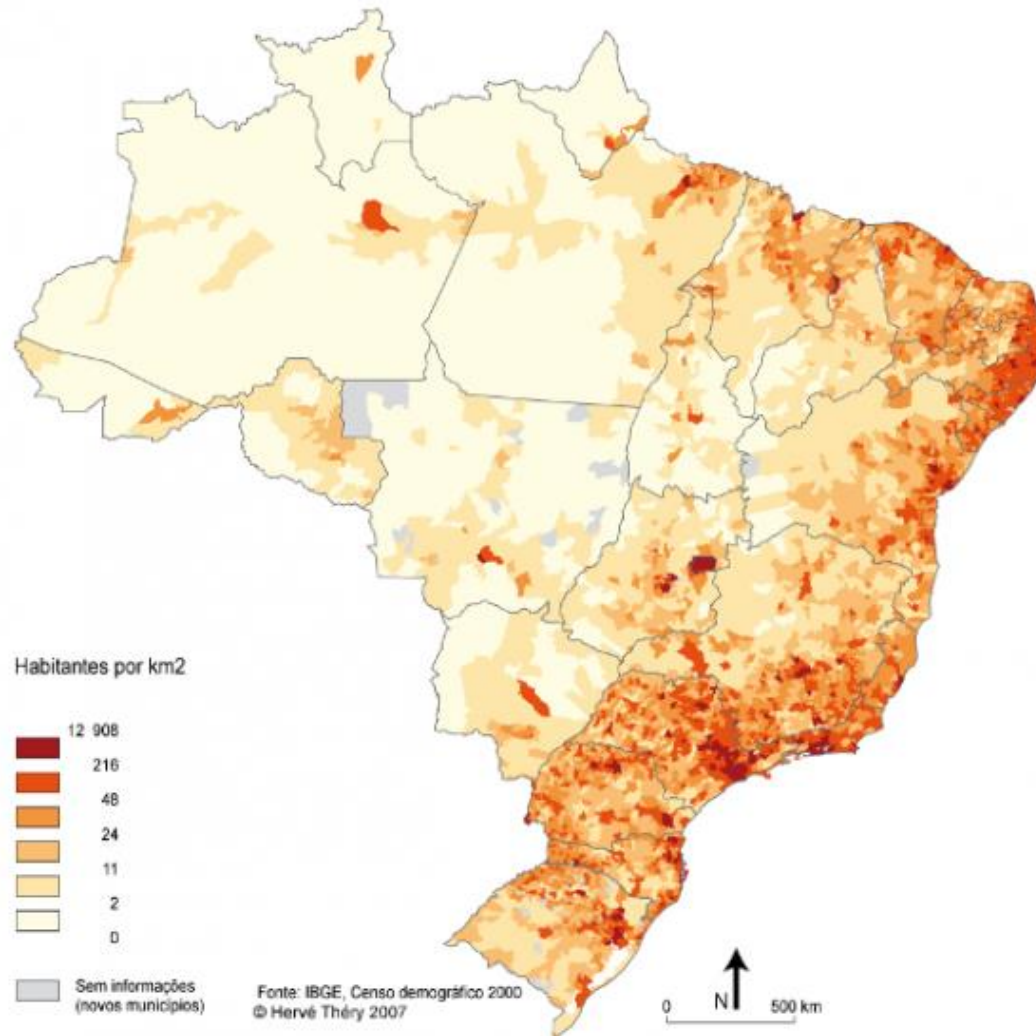
- Para que a tabela seja facilmente compreendida é necessário:
 - Simplicidade e clareza
 - Que seja autoexplicativa
 - Possuir título identificando
 - Explicar as variáveis e suas unidades
 - Incluir rodapé quando necessário
 - Citar a Fonte



Apresentação gráfica

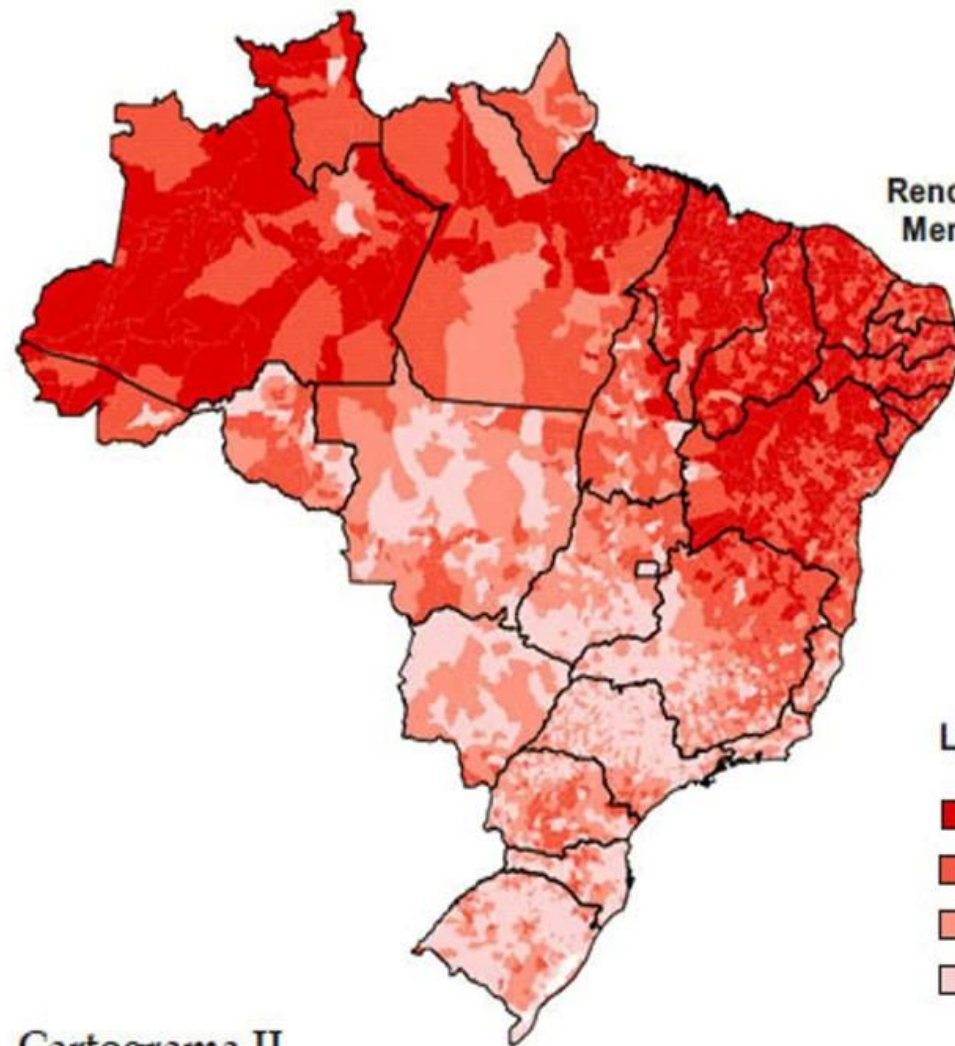
- A apresentação gráfica é um complemento importante da apresentação tabular
- A principal vantagem de um gráfico sobre a tabela é o fato de que ele permite uma visualização imediata da distribuição dos valores observados

Densidade de povoamento



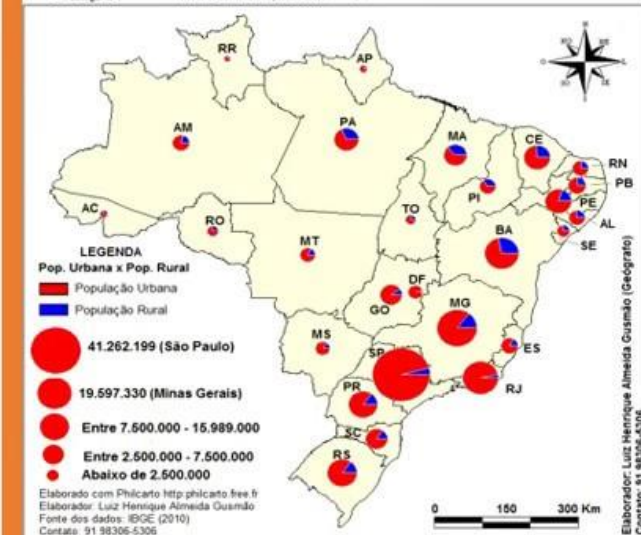
Tipos de gráficos

- Cartogramas - são mapas geográficos em que as frequências das categorias de uma variável são projetadas nas áreas específicas do mapa.
- Os Cartogramas são amplamente difundidos e utilizados em Geografia, História e Demografia.
- Os cartogramas são indicados para representar séries geográficas.



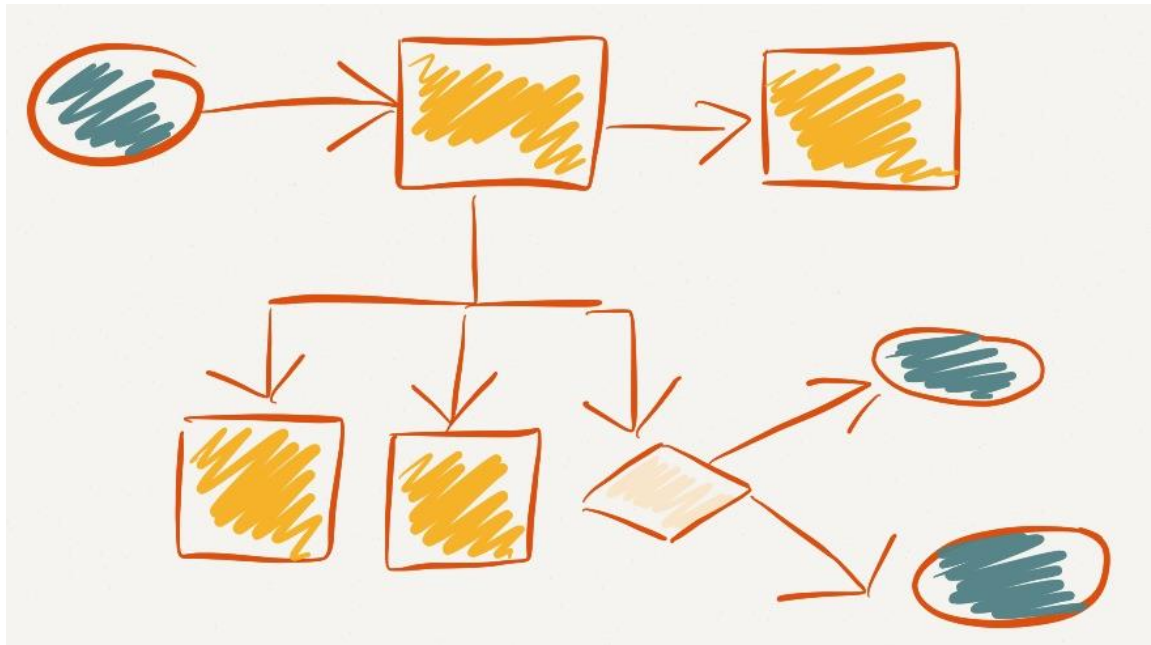
Cartograma II
Elaboração: CGMA/SDR/MI

Fluxos de passageiros

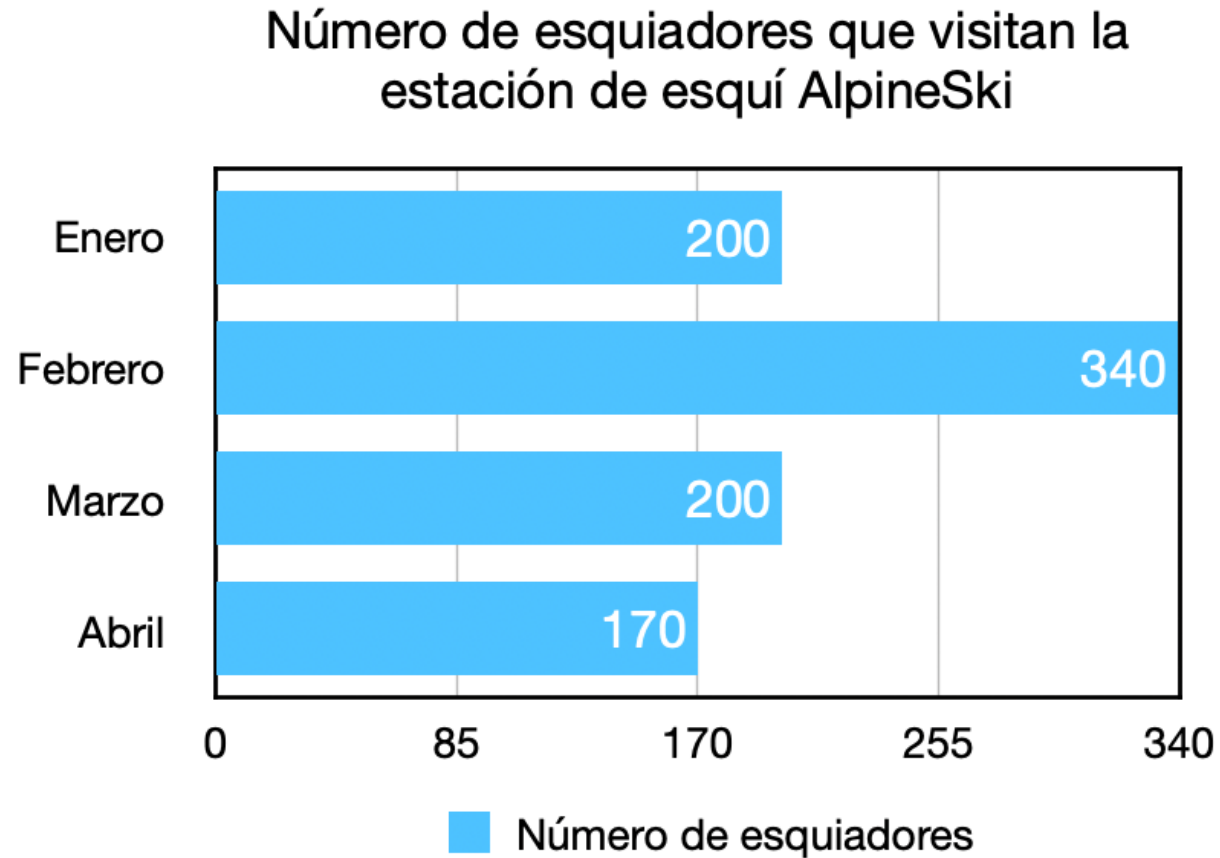


Tipos de gráficos

- Diagramas - são gráficos em que a magnitude das frequências é representada por certa mensuração de uma determinada figura geométrica.
- São os gráficos mais usados na representação de séries estatísticas e se apresentam através de uma grande variedade de tipos.

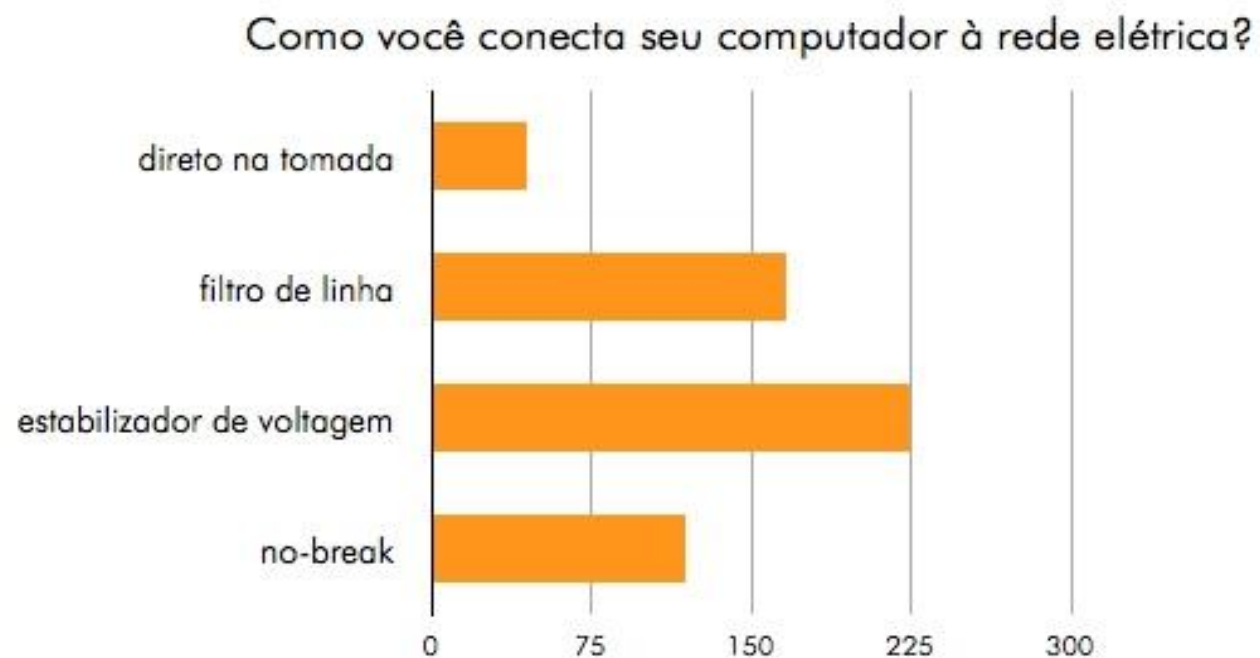


Tipos de gráficos



- Gráfico de Barras e Gráfico de Colunas
- São indicados para representar séries categóricas (variáveis qualitativas) e séries geográficas.

Gráfico de barras



Fonte: Pesquisa Zumo de No-Breaks, maio de 2009.

Ocorrências por Tipo de Deficiência

Fonte: Censo 2000

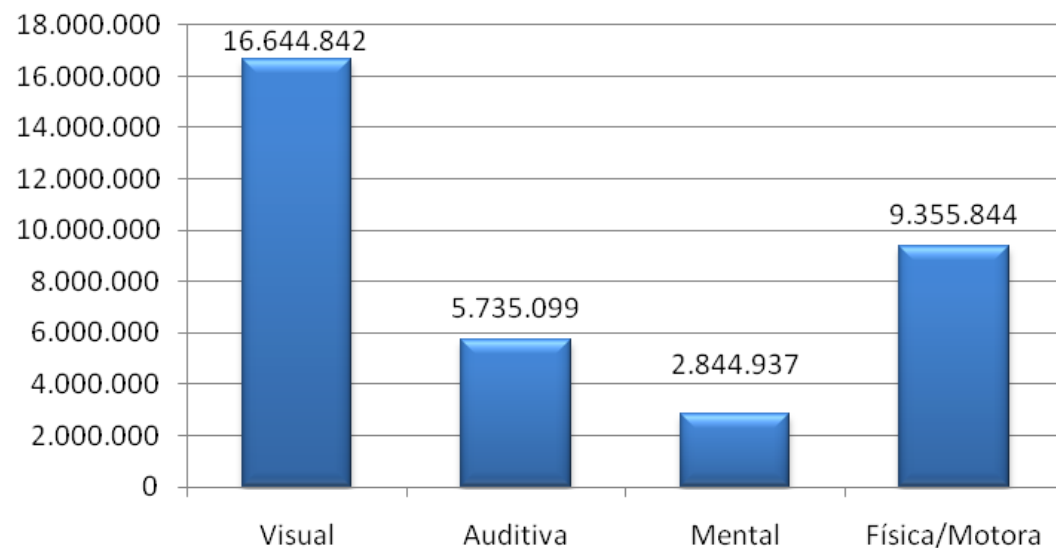


Gráfico de colunas

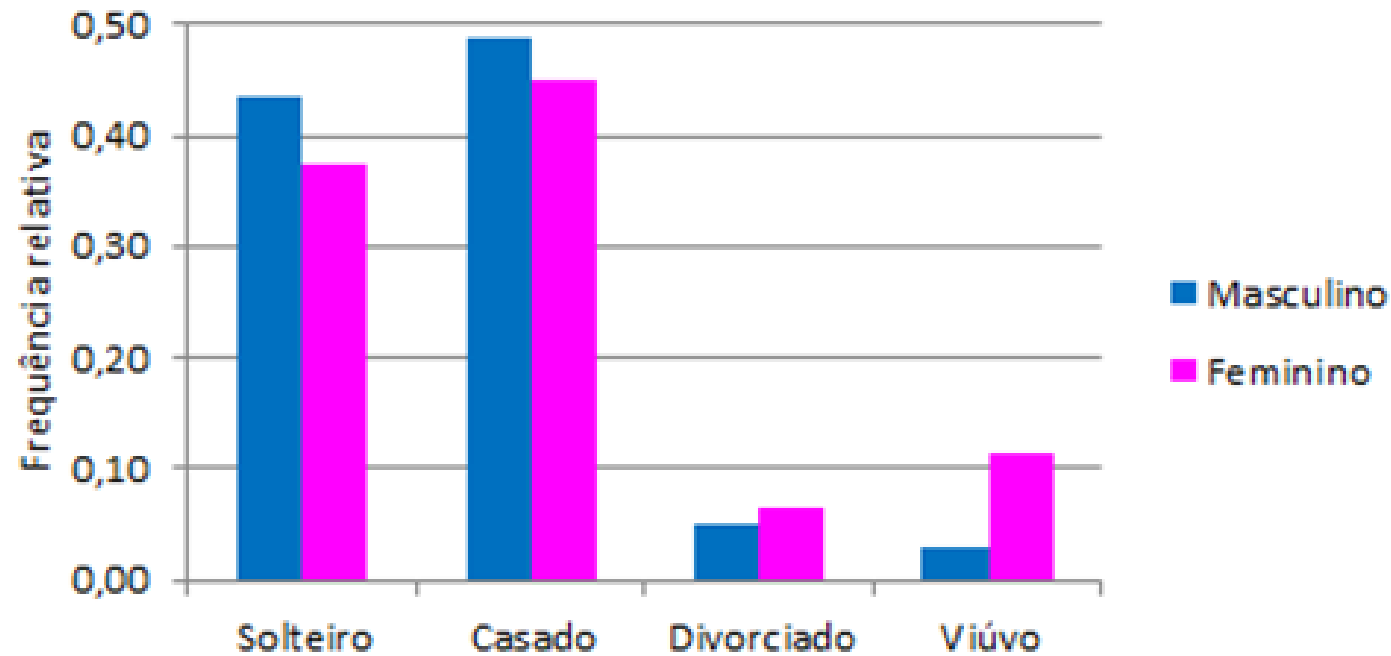
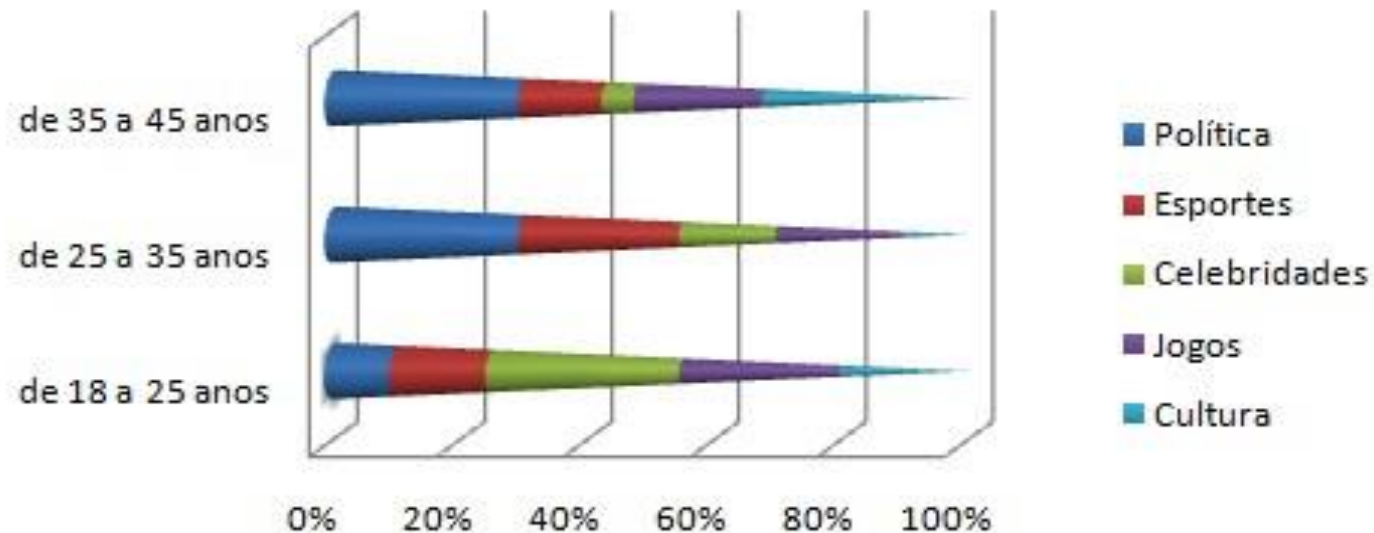


Gráfico de barras múltiplas

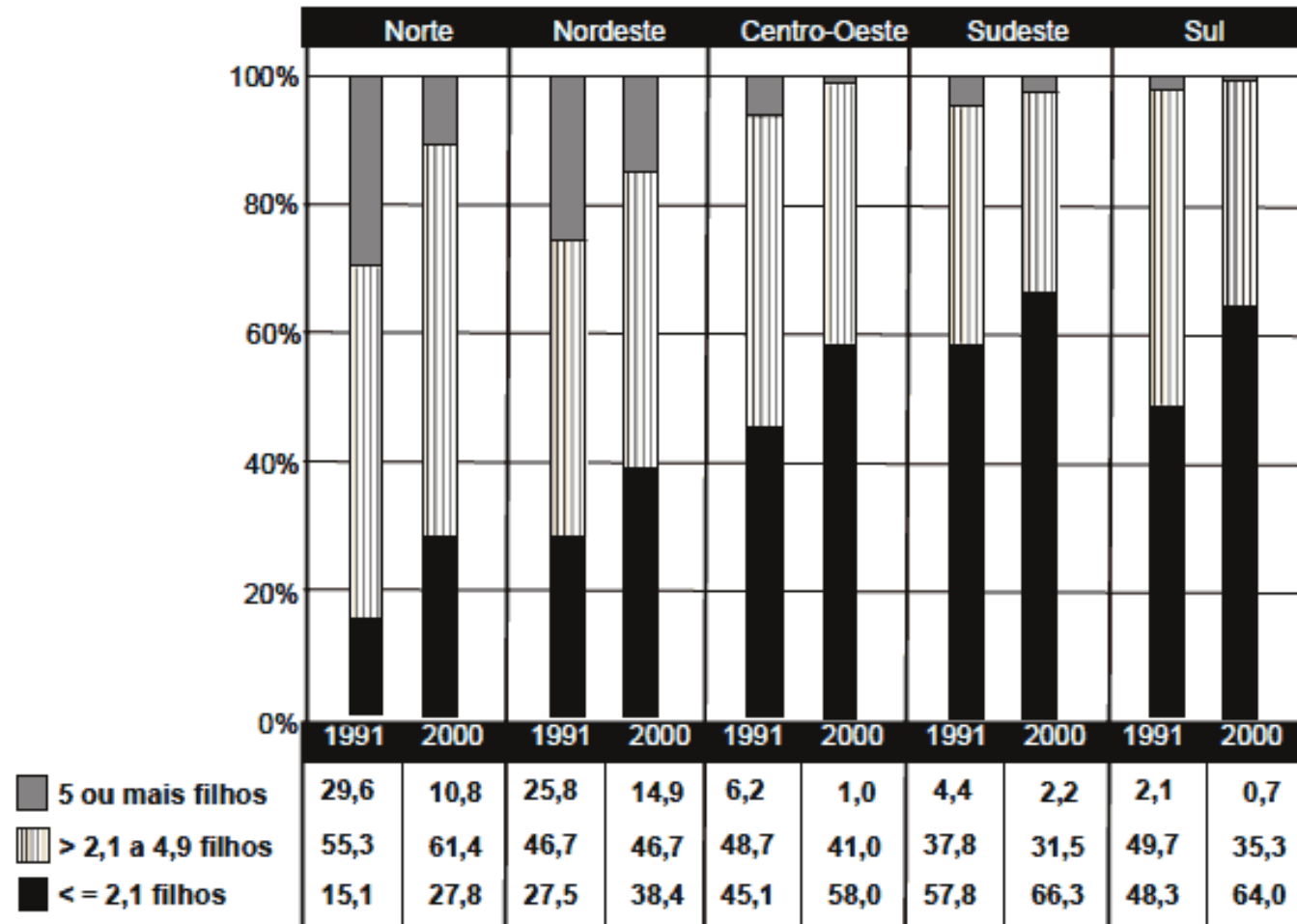
Gráfico de barras empilhadas

Usuários por faixa etária versus conteúdo

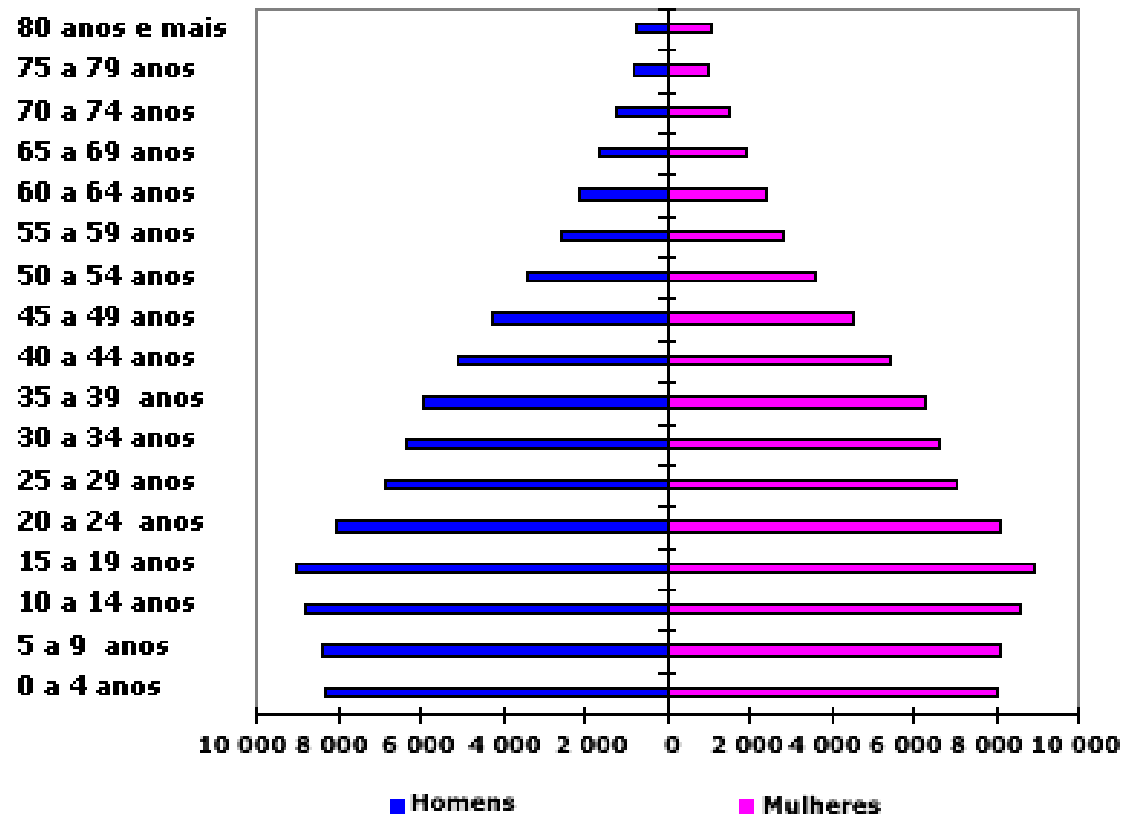


- Utilizado para ilustrar uma representação proporcional dentro de um conjunto de dados

Gráfico de barras empilhadas



Pirâmide etária absoluta - 2000

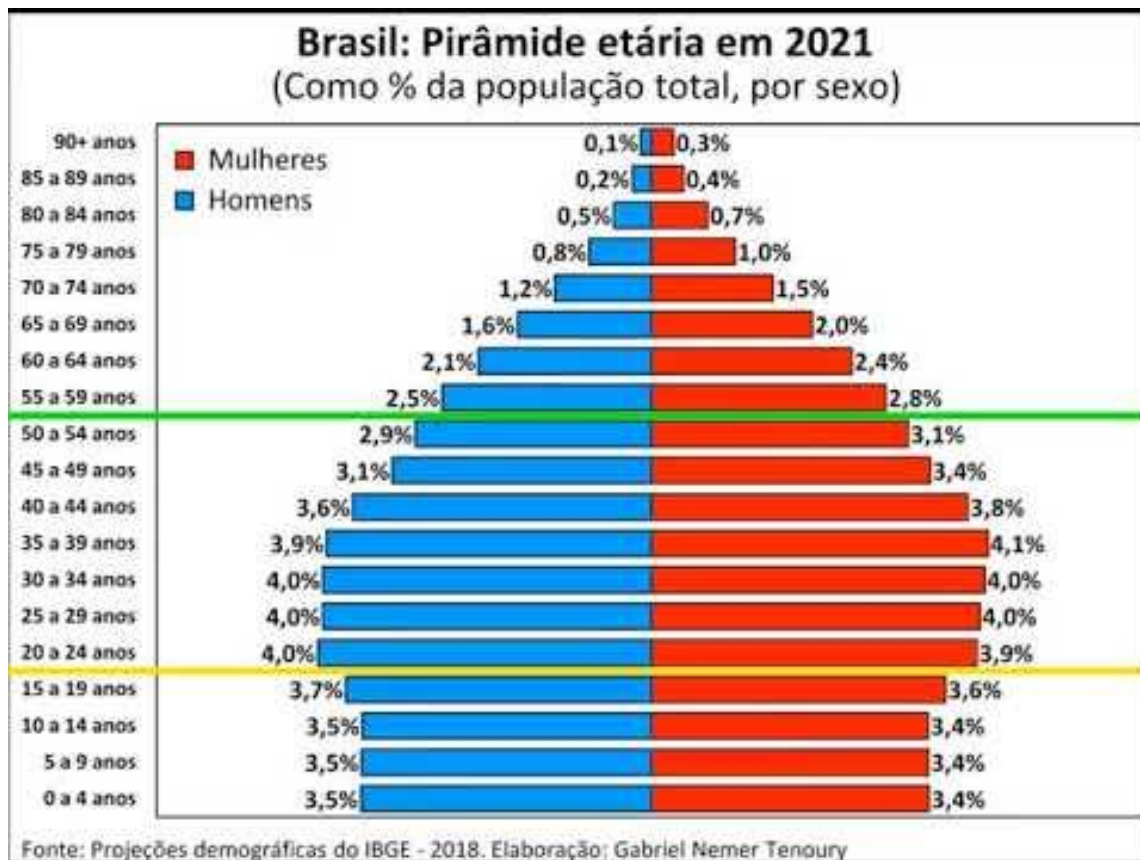


Fonte: IBGE, Censo Demográfico 2000.

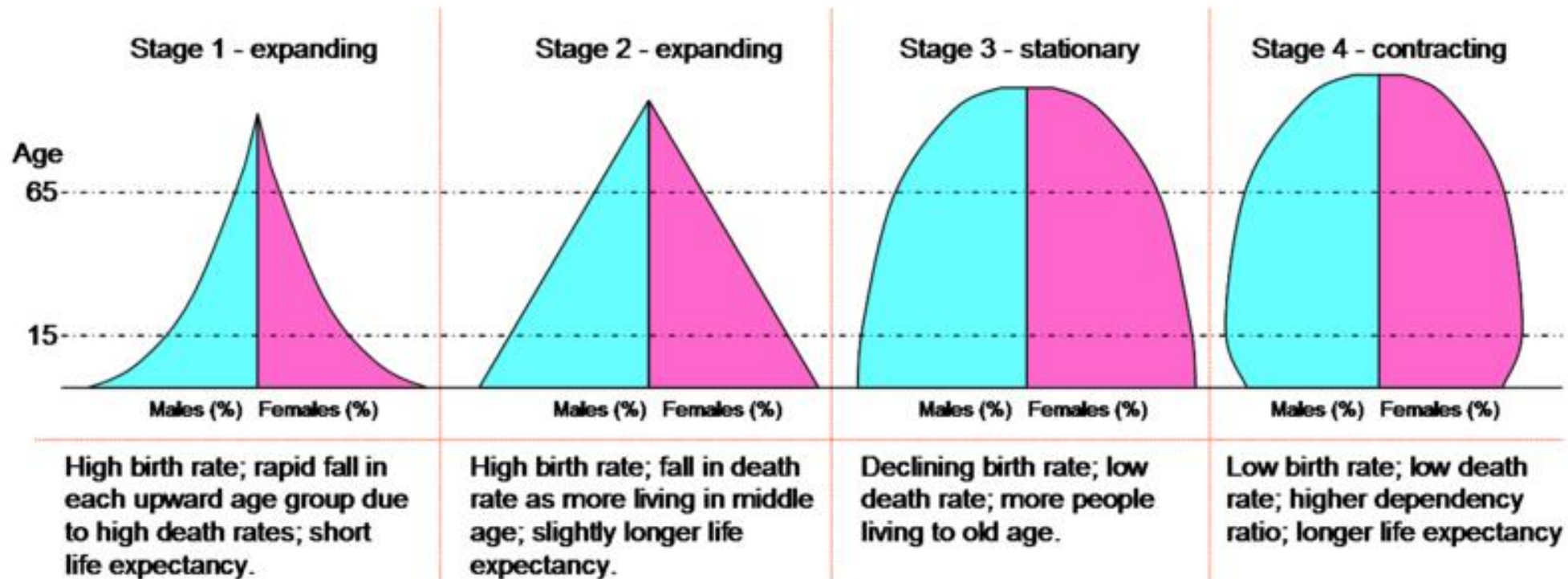
Pirâmide Etária

- Também conhecida como pirâmide demográfica ou pirâmide populacional.
- Mostra a distribuição de diferentes grupos etários em uma população, em que normalmente cria-se a forma de uma pirâmide.
- Esse gráfico é constituído de dois conjuntos de barras que representam o sexo e a idade.

Pirâmide Etária

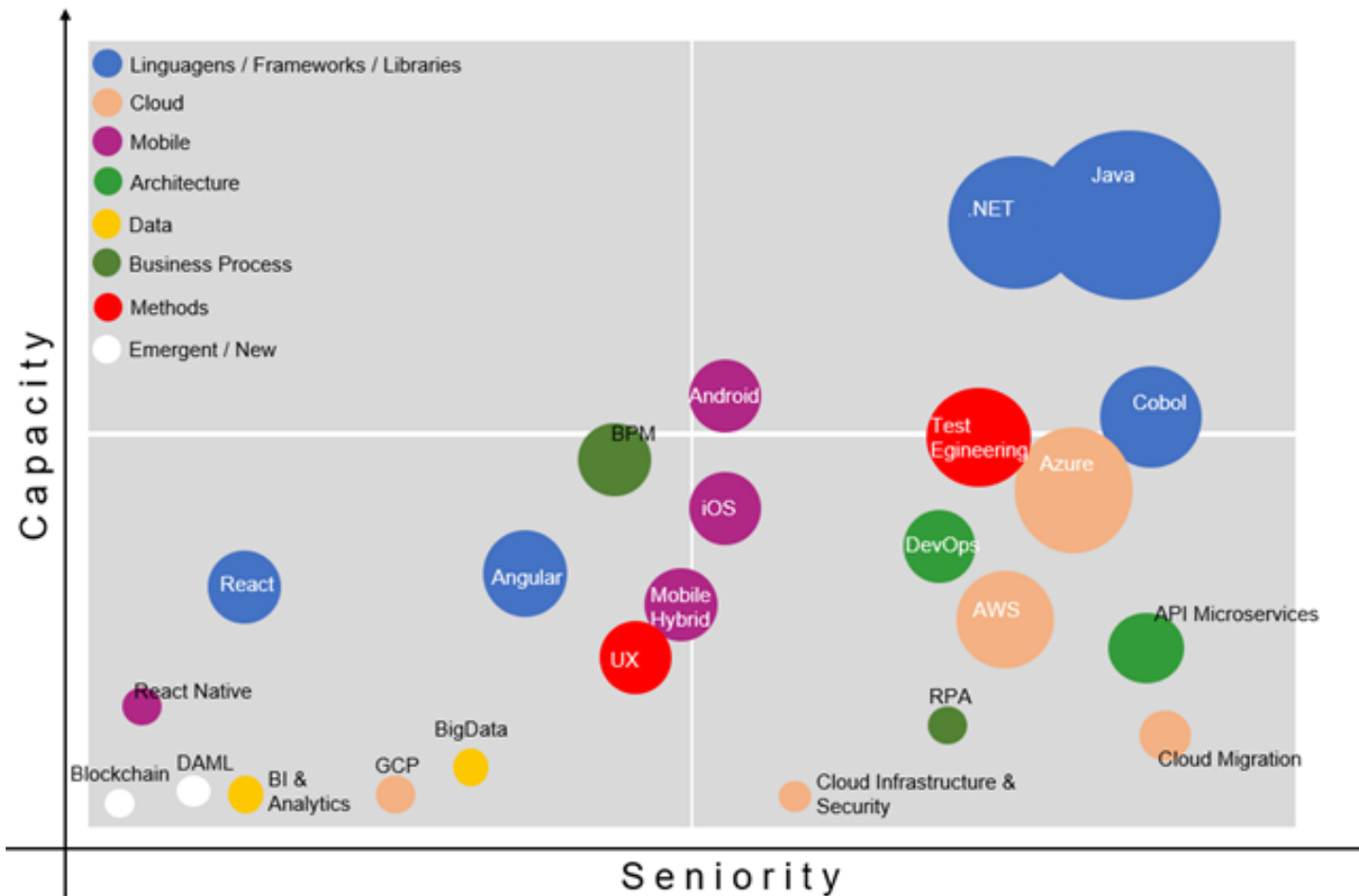


- São usadas, não só para monitorar a estrutura de sexo e idade, mas como um complemento aos estudos da qualidade de vida, já que podemos visualizar a média do tempo de vida, a taxa de mortalidade e a regularidade, ou não, da população ao longo do tempo.
- Quanto mais alta a pirâmide, maior a expectativa de vida e, consequentemente, melhor as condições de vida daquela população.
- É possível perceber que quanto mais desenvolvido economicamente e socialmente é o país, mais sua pirâmide terá uma forma retangular.



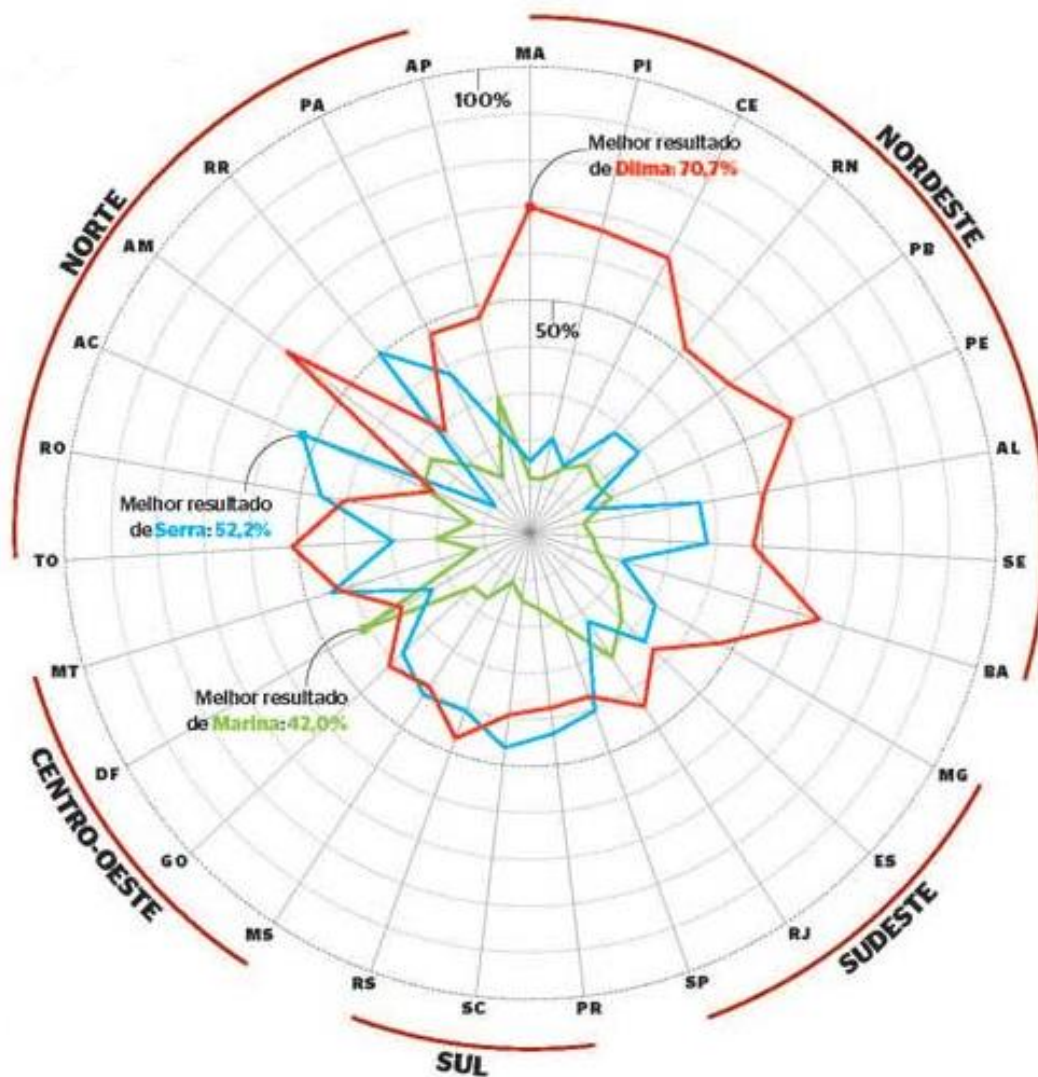
Piramide Etária

Gráfico de bolhas



- Podem ser plotados em um gráfico de bolha dados que estão organizados em colunas em uma planilha, de tal forma que valores de X sejam listados na primeira coluna, enquanto valores de Y correspondentes e valores de tamanho de bolha estejam listados em colunas adjacentes.

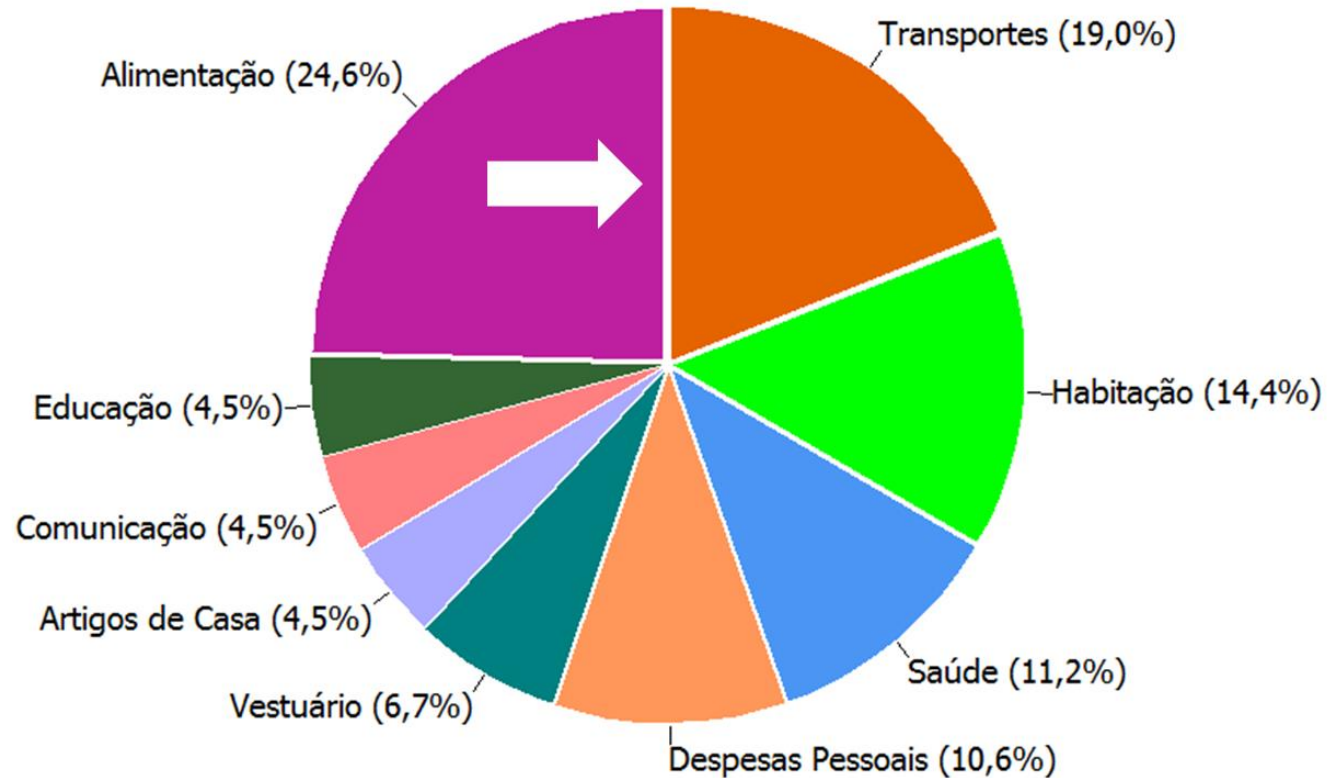
Gráfico de radar



- Os dados organizados em colunas ou linhas em uma planilha podem ser plotados em um gráfico de radar.
- Gráficos de radar comparam os valores agregados de várias séries de dados.

Gráfico de setores ou pizza

Peso dos itens no IPCA (em jan/14)



- Utilizado para representar uma série categórica (variáveis qualitativas)
- Neste tipo de gráfico todo o conjunto de dados é representado por um círculo, e cada categoria é representada por parte desse círculo (isto é, um setor).

Vendas Trimestrais

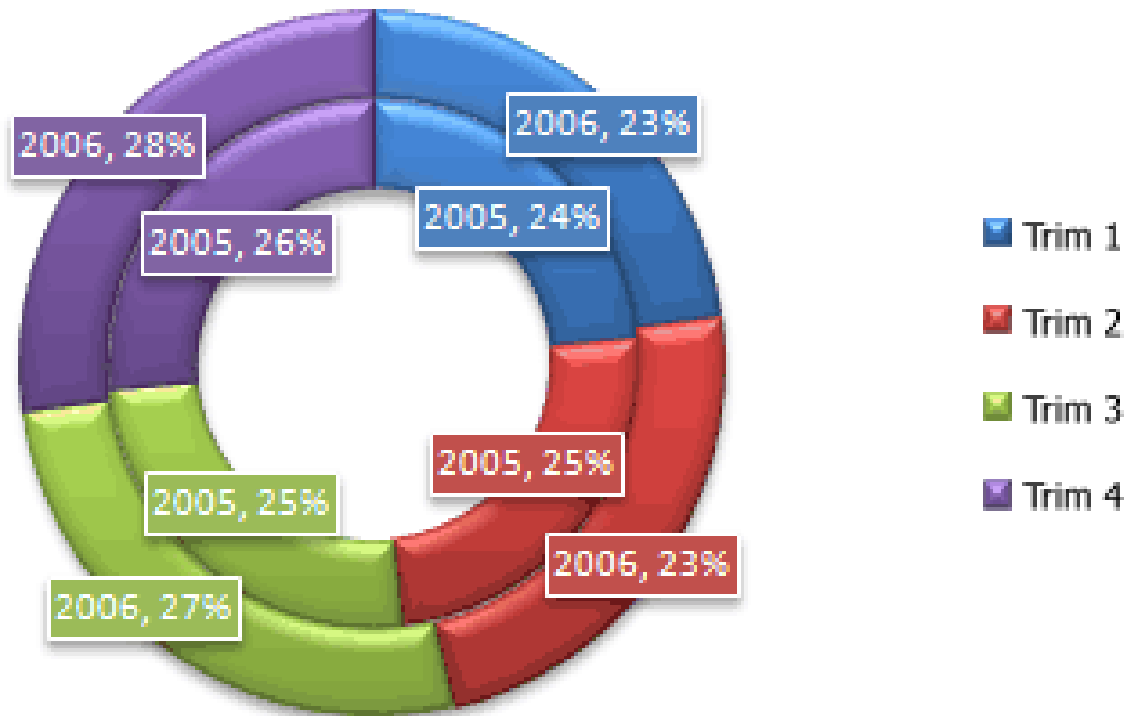
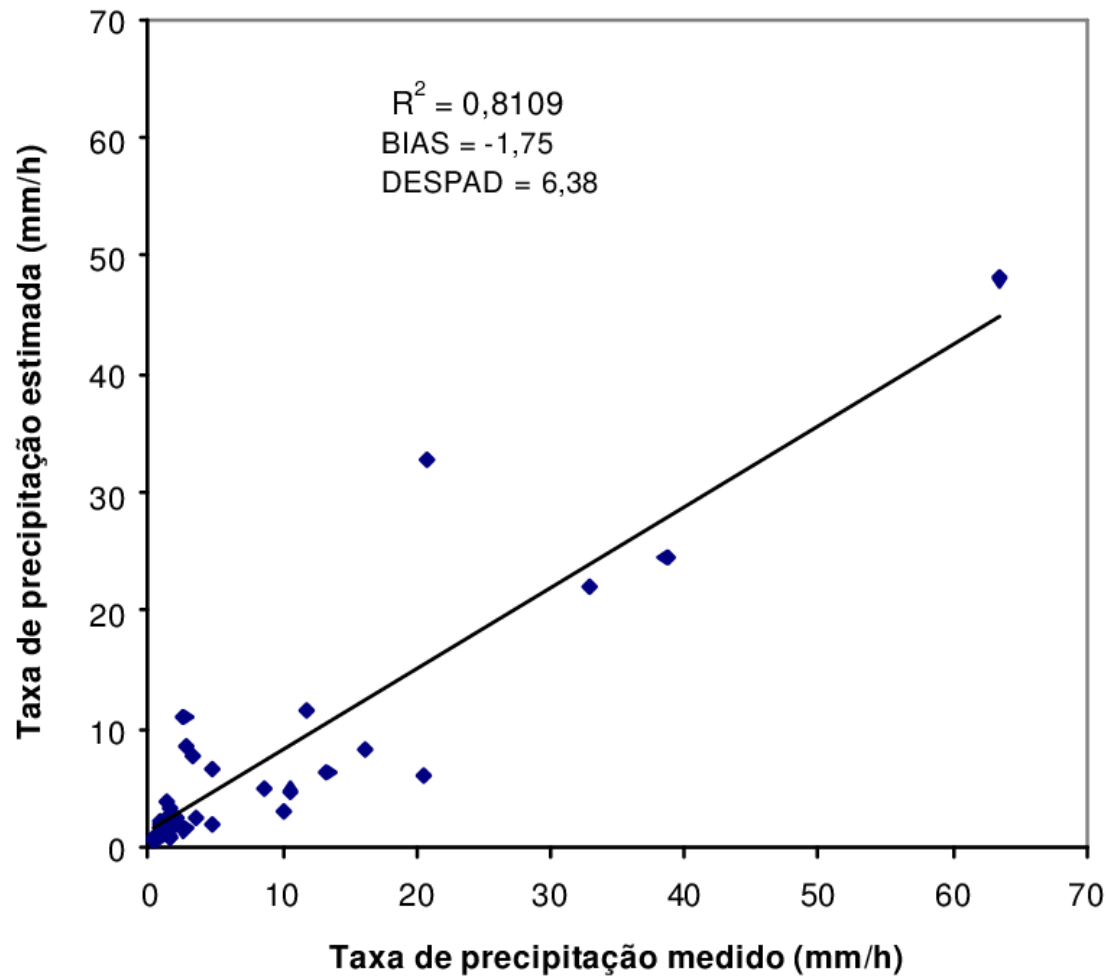


Gráfico de rosca

Gráfico de dispersão



- Gráficos de dispersão mostram as relações entre os valores numéricos em várias sequências de dados ou plotam dois grupos de números como uma sequência de coordenadas XY.

MOVIMENTO DE PROCESSOS

Tribunais judiciais de primeira instância,
balanço dos últimos dez anos



Gráfico de linhas

- Representa observações feitas ao longo do tempo, em intervalos iguais ou não.

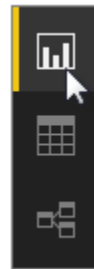


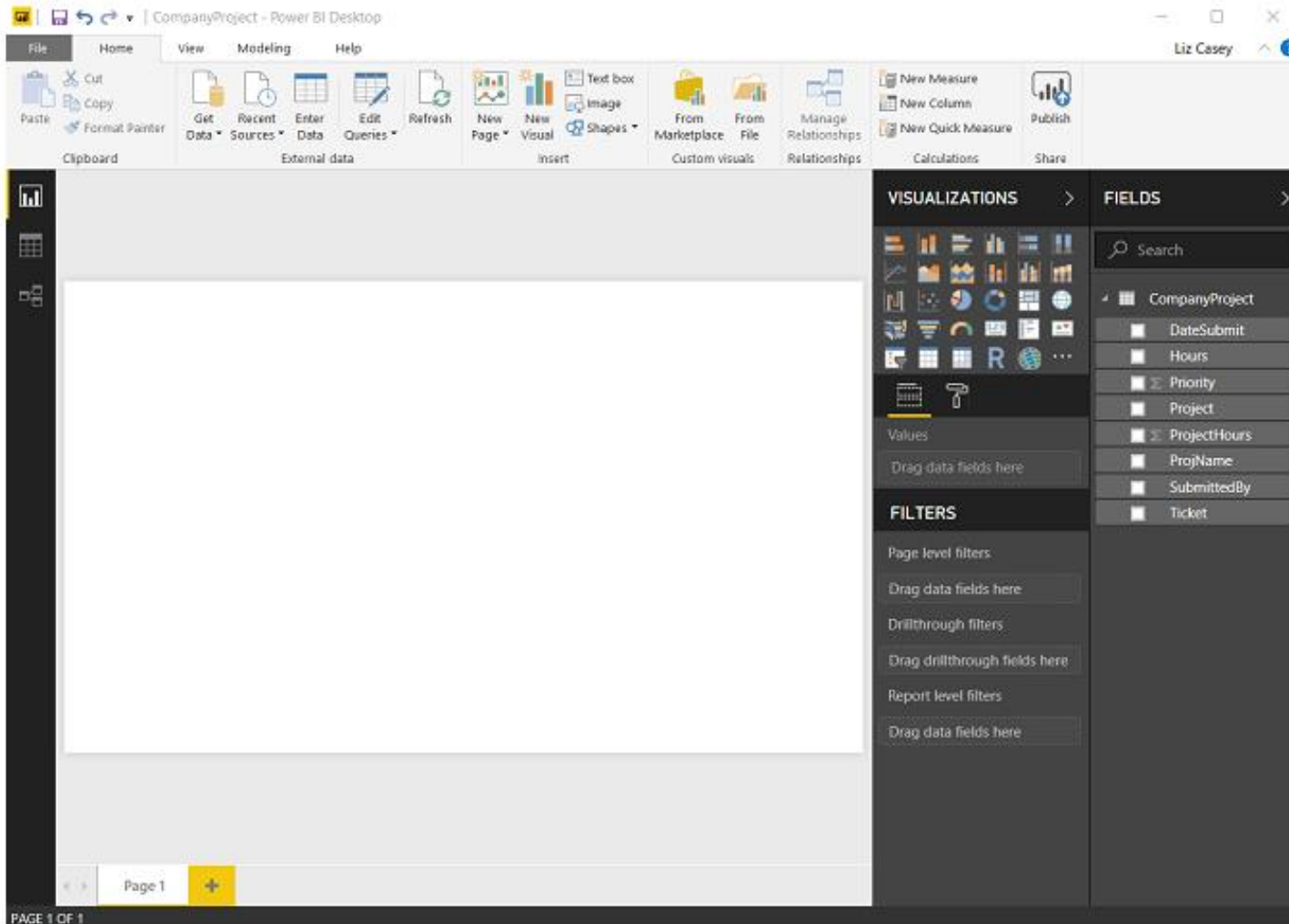
Pictogramas

- São construídos a partir de figuras ou conjuntos de figuras representativas da intensidade ou das modalidades do fenômeno.

Criar relatórios no Power BI

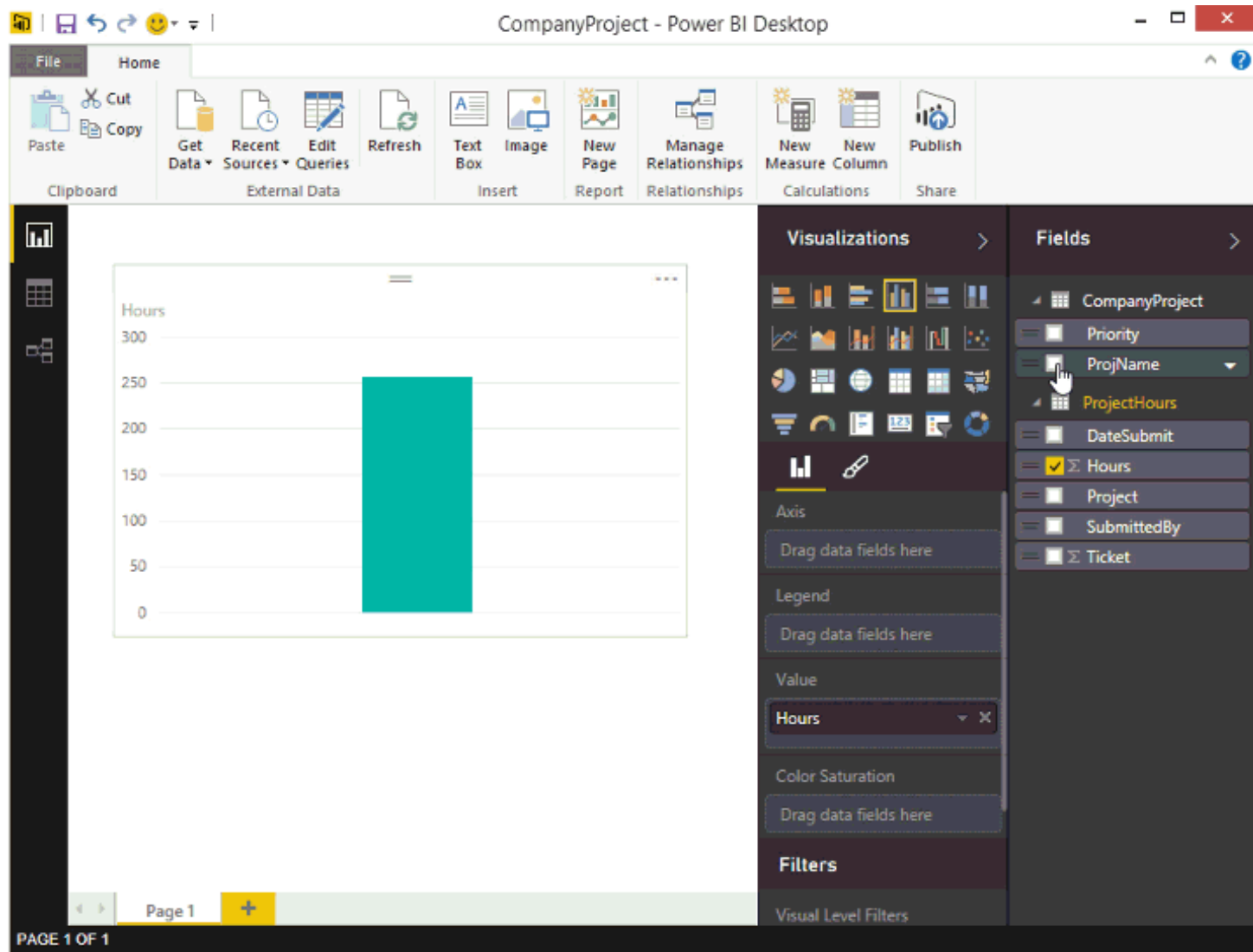
- No Power BI o “Modo de Exibição de Relatório”, é a área onde os relatórios são criados e manipulados
- Um relatório pode ter quantas páginas desejar, e os objetos podem ser movidos de um lugar para outro.
- De acordo com o site da Microsoft, você pode alternar entre o “Modo de Exibição de Relatório”, o “Modo de Exibição de Dados” e o “Modo de Exibição de Relação” selecionando os ícones na barra de navegação à esquerda:





Criar relatórios no Power BI

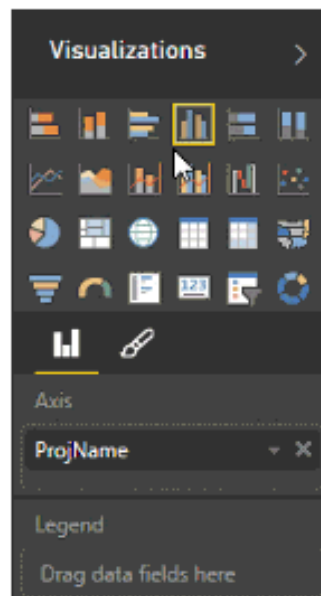
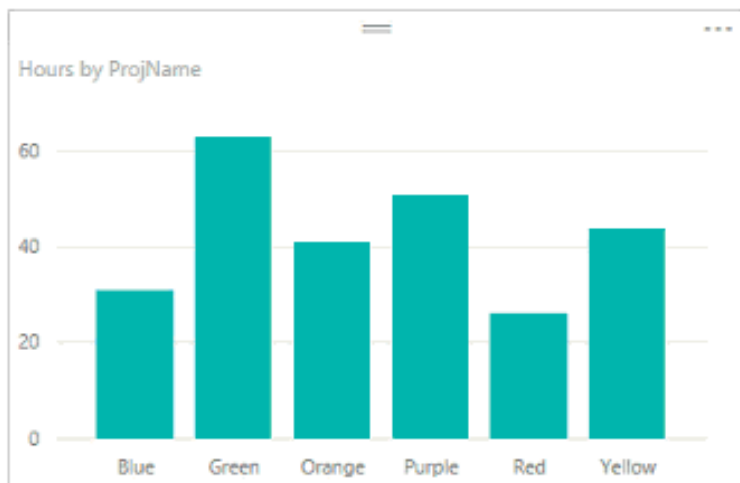
- Após carregar dados pela primeira vez no Power BI Desktop, você verá o “Modo de Exibição de Relatório” com uma tela em branco.



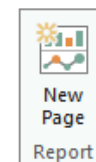
Criar relatórios no Power BI

- Na sequência, você pode adicionar campos a uma nova visualização na tela. Conforme imagem:

Criar relatórios no Power BI

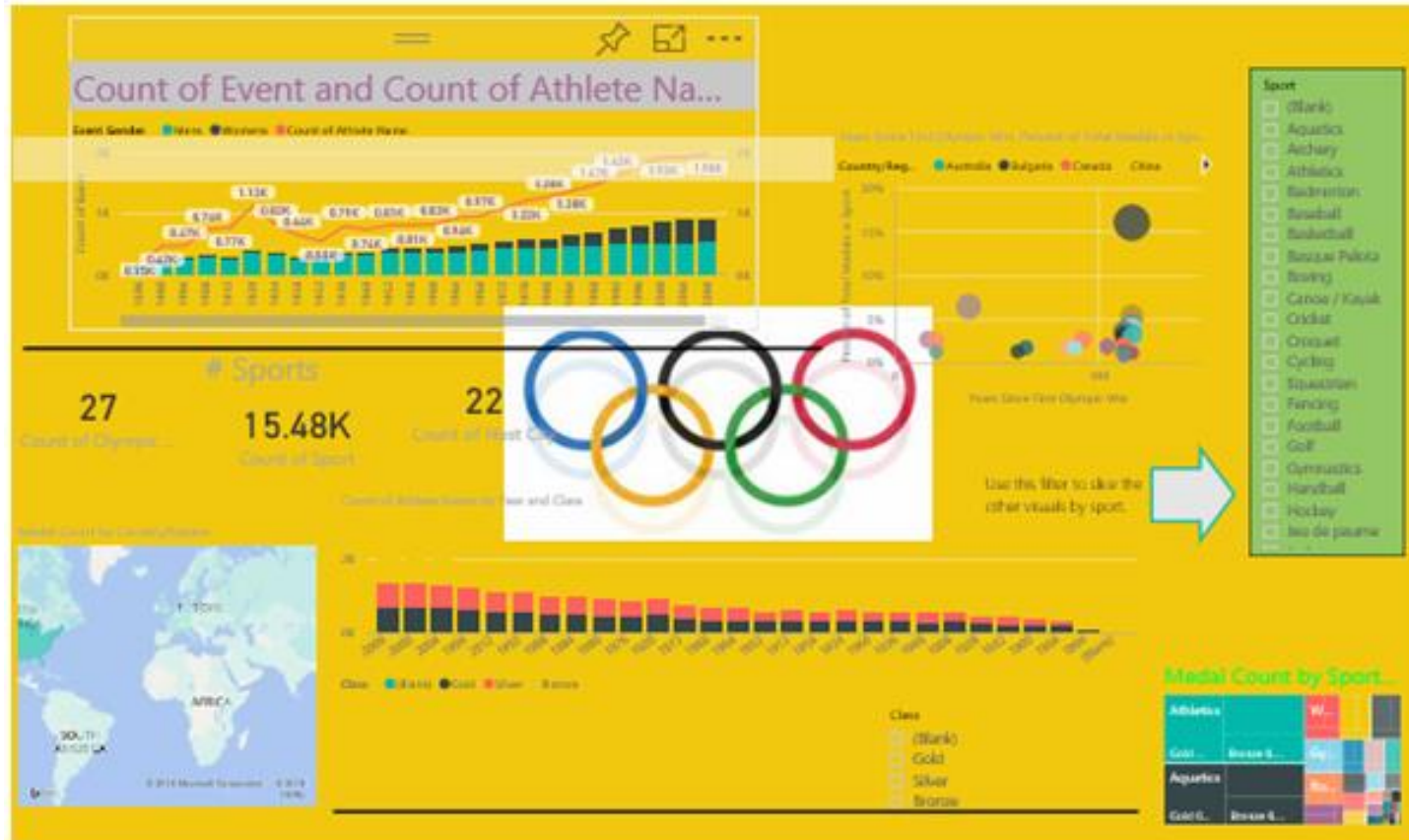


- Para alterar o tipo de visualização, você poderá selecioná-la no grupo Visualizar na faixa de opções, ou então você pode clicar com o botão direito do mouse e selecionar um tipo diferente no ícone “Alterar tipo de visualização”
- Você pode adicionar novas páginas ao seu relatório. Basta clicar em “Nova Página” na faixa de opções.

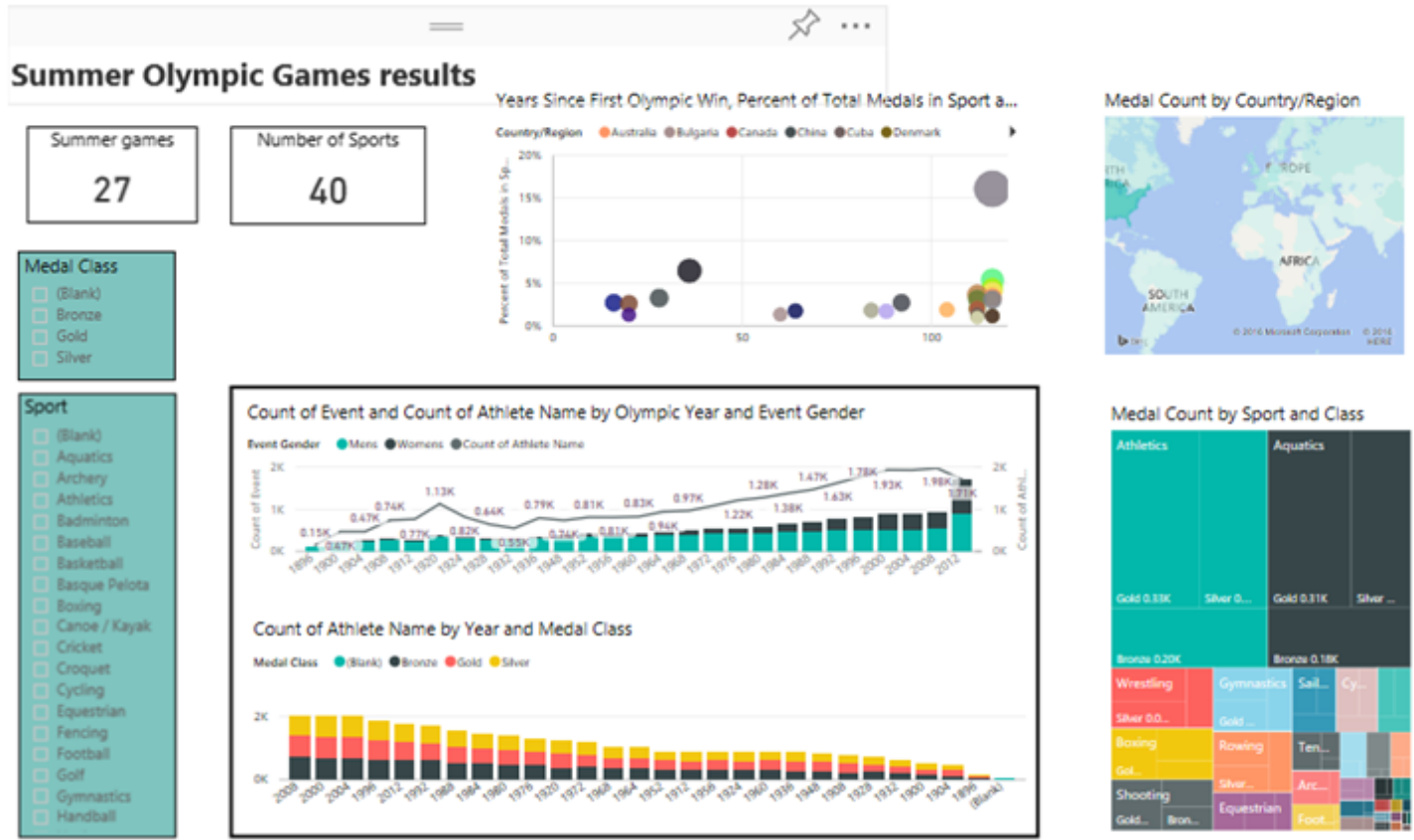


Criar relatórios no Power BI

- Relatório ANTES de aplicar melhores práticas de visualização de dados



Criar relatórios no Power BI

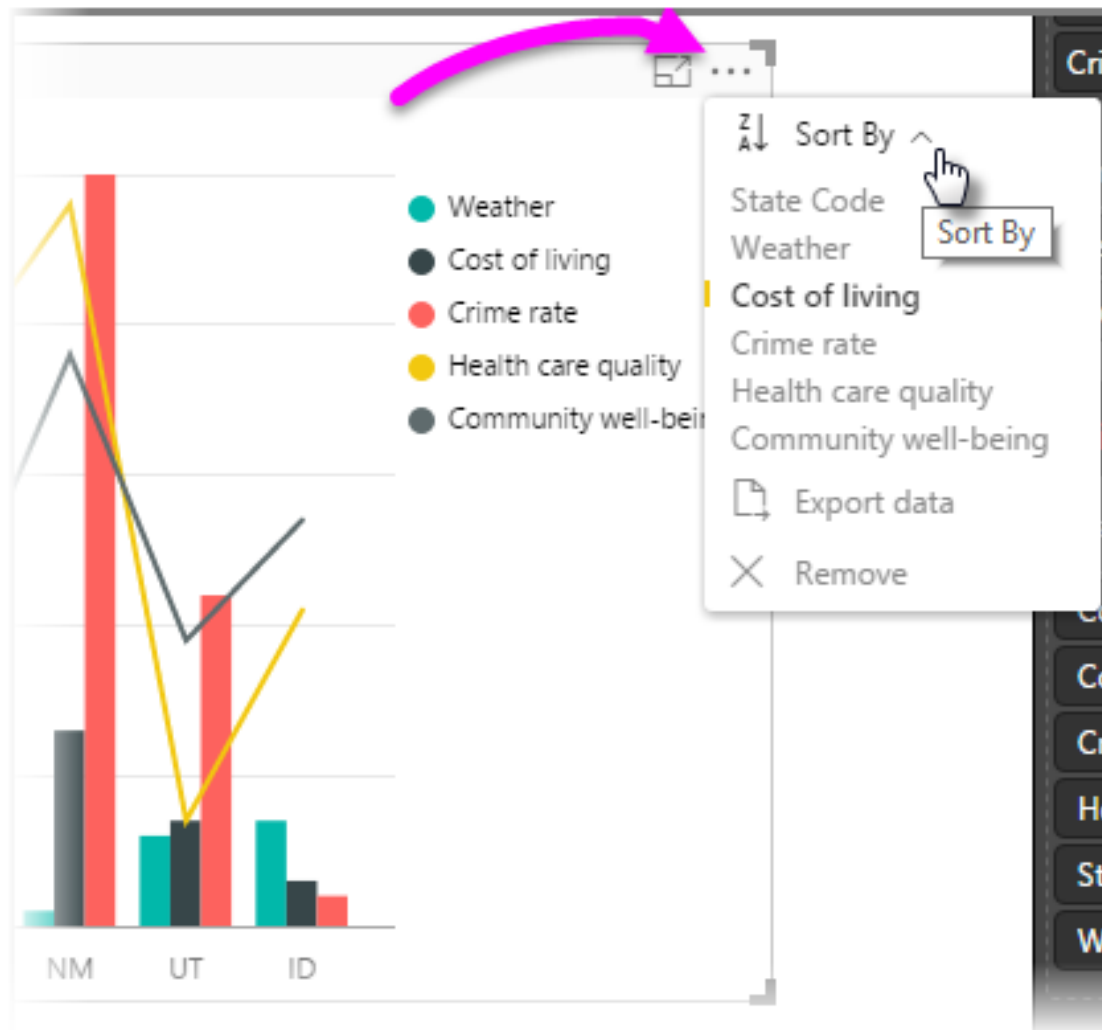


- Relatório DEPOIS de aplicar melhores práticas de visualização de dados

Classificação no Power BI

- No Power BI você pode alterar a aparência de uma visualização, classificando-a por diferentes campos de dados.
- Ao alterar a maneira como você classifica uma visualização, é possível realçar as informações que você deseja transmitir e garantir que a visualização reflita qualquer tendência (ou ênfase) que você deseja transmitir.

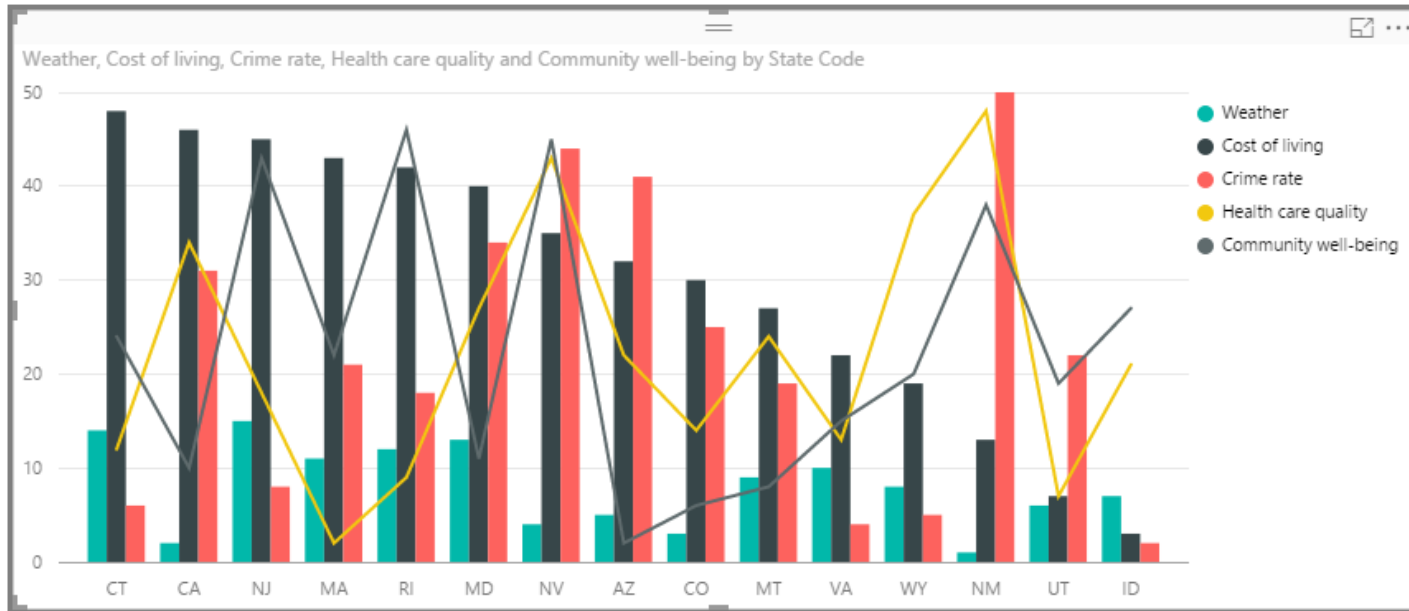




Classificação no Power BI

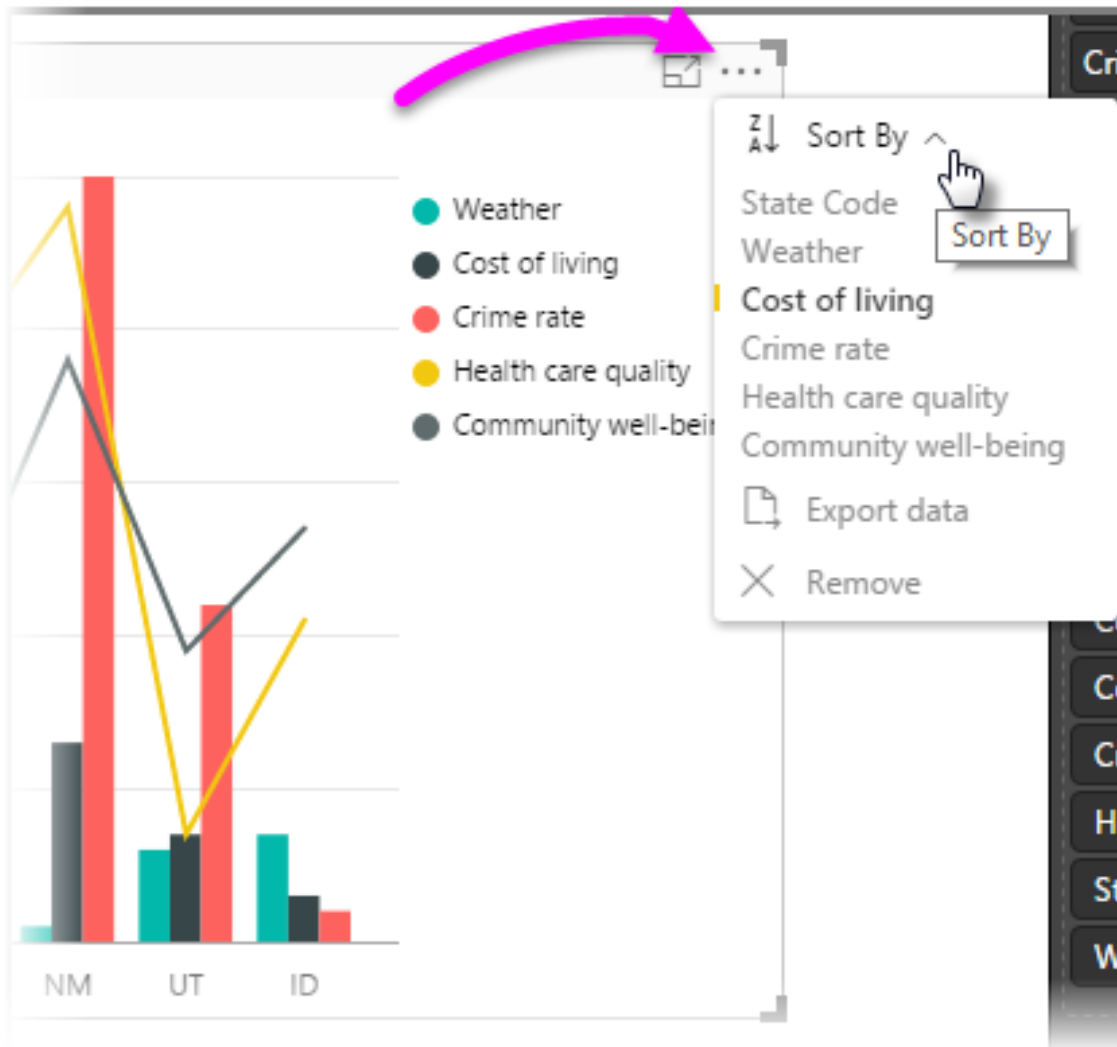
- Se estiver usando dados numéricos (como valores de vendas) ou dados de texto (como nomes de estado), você pode classificar suas visualizações da forma que quiser e fazer com que elas tenham a aparência que você quiser.
- O Power BI oferece muita flexibilidade para a classificação e menus rápidos para você usar.
- Em qualquer visual, selecione o menu de reticências (...) e, em seguida, Classificar por.
- Em seguida, selecione o campo pelo qual você deseja classificar, conforme mostrado na imagem.

Classificação no Power BI

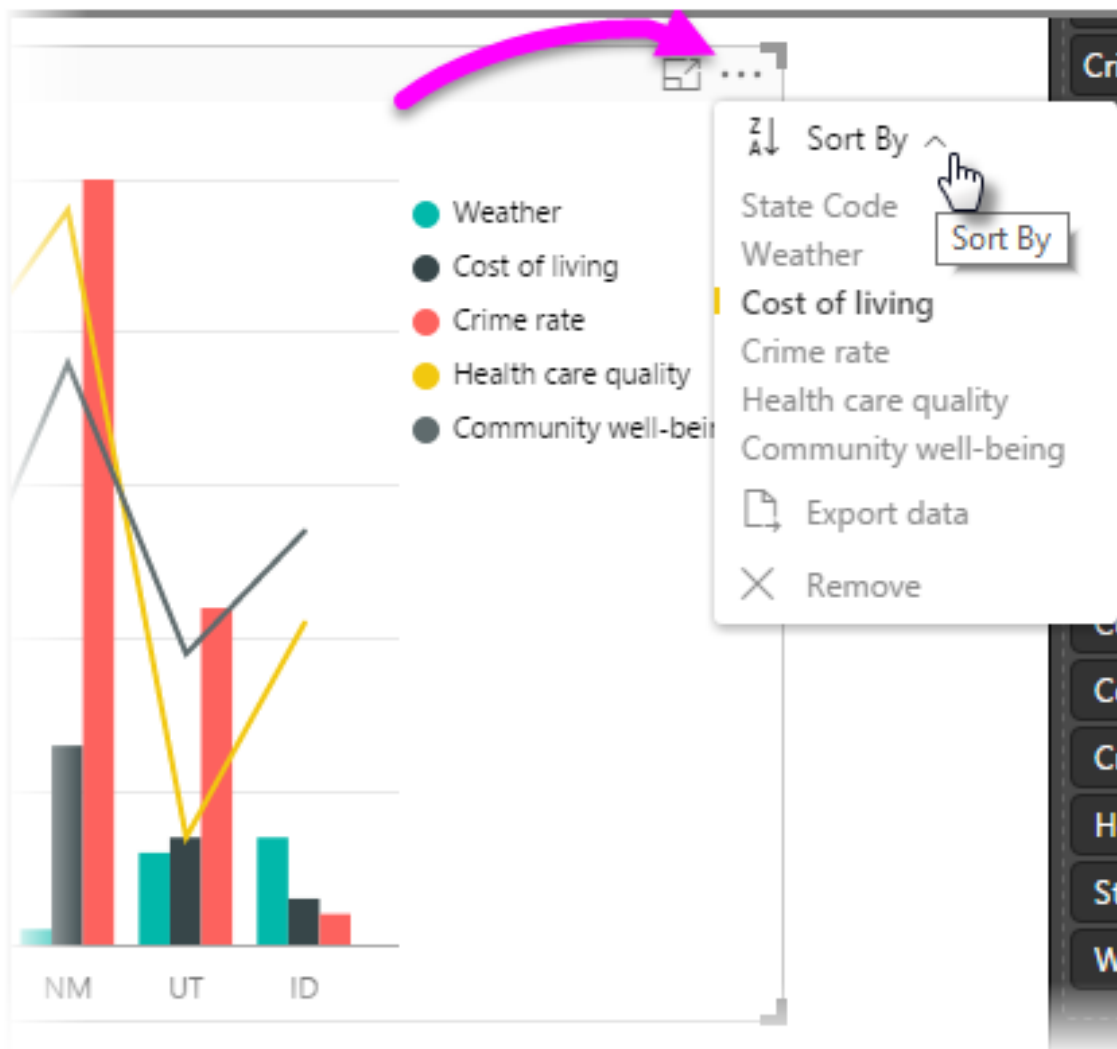


- Vamos ver com mais detalhes como a classificação funciona no Power BI.
- A visualização a seguir lista os 15 primeiros estados em termos de clima (dias mais ensolarados, classificados de 1 a 50, com 1 tendo mais dias ensolarados).
- Esta é a visualização como ela aparece antes de fazermos qualquer classificação.

Classificação no Power BI



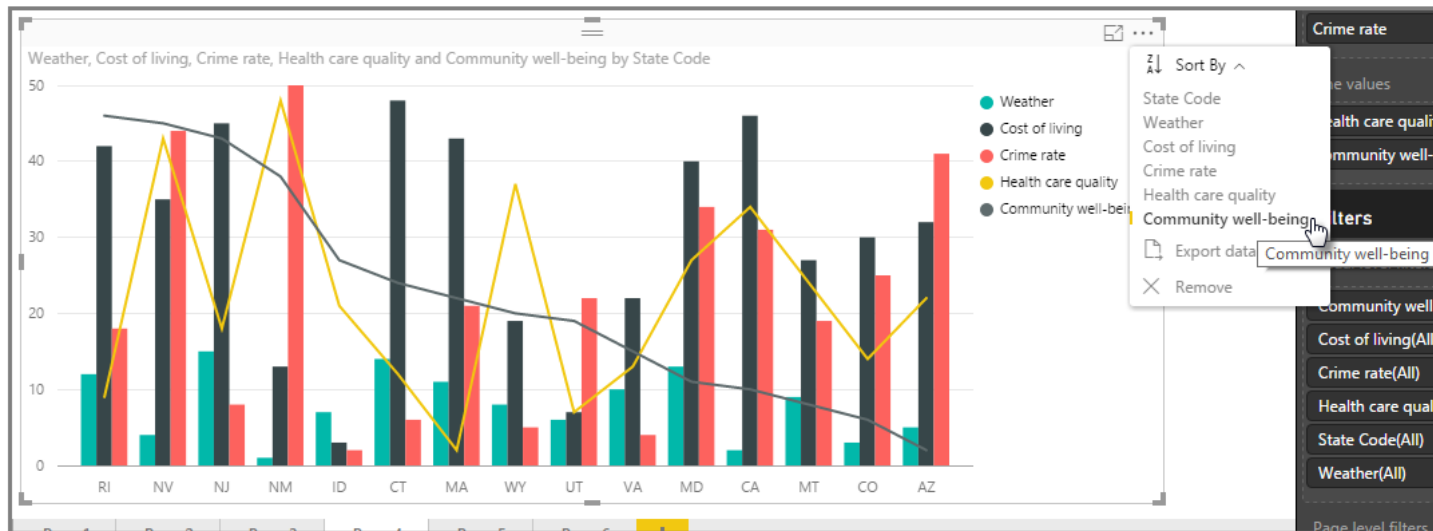
- Nessa imagem a visualização está classificada por Custo de vida – podemos observar isso fazendo a correspondência entre a cor das barras decrescentes e a legenda, mas há uma maneira melhor de determinar a coluna de classificação atual:
 - o diálogo Classificar Por, disponível no menu de reticências (...) no canto superior direito do visual.
- Quando selecionamos as reticências, vemos o seguinte, na imagem.



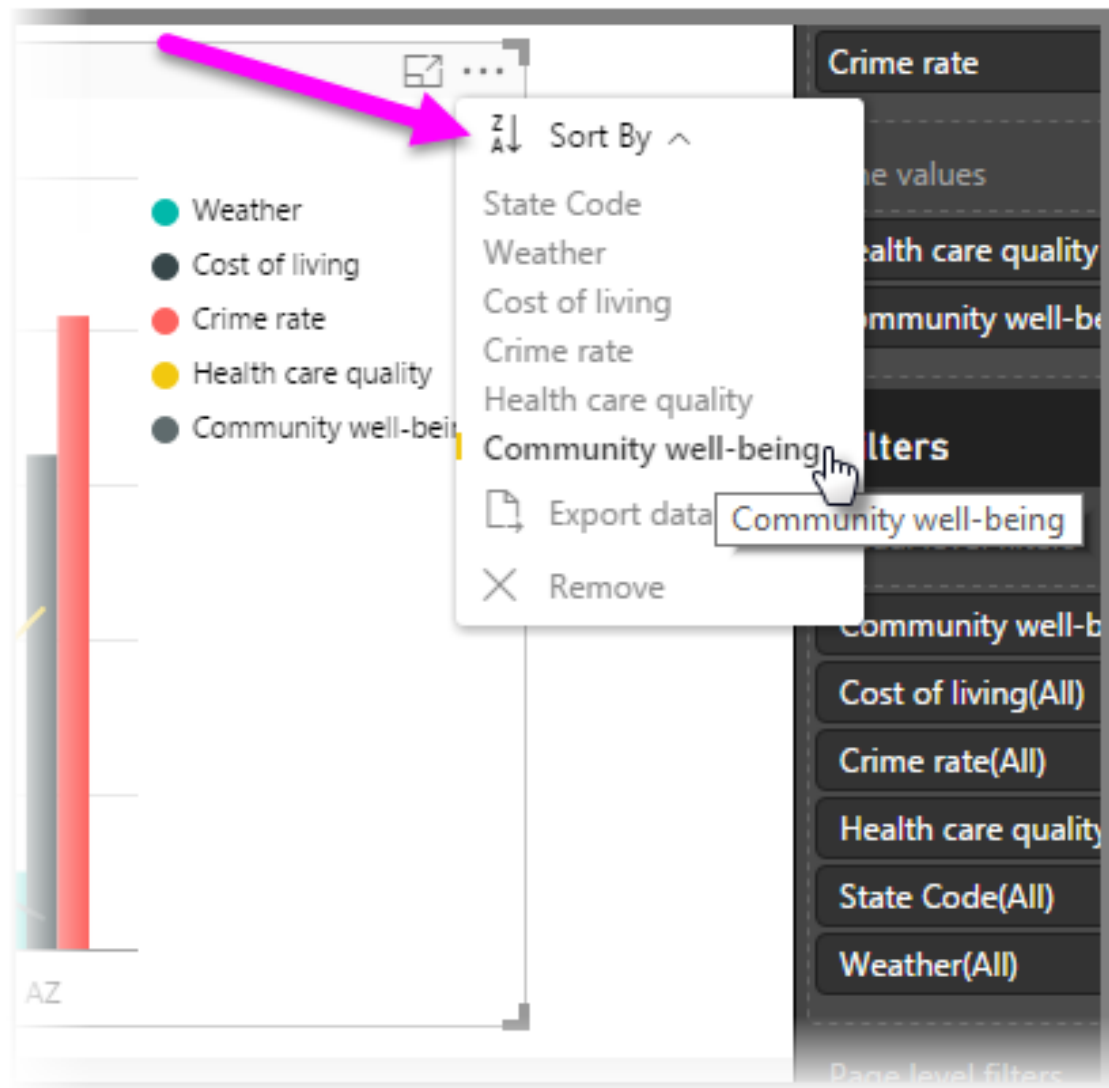
Classificação no Power BI

- Há alguns itens a serem observados no menu que aparece quando você seleciona as reticências:
- Na barra amarela ao lado de Custo de vida e o fato de Custo de vida estar em negrito;
- O pequeno ícone ao lado das palavras Classificar Por, que mostra Z/A (Z sobre A) e uma seta para baixo.
- Vamos examinar cada um deles independentemente.

Selecionando a coluna a ser usada para classificação

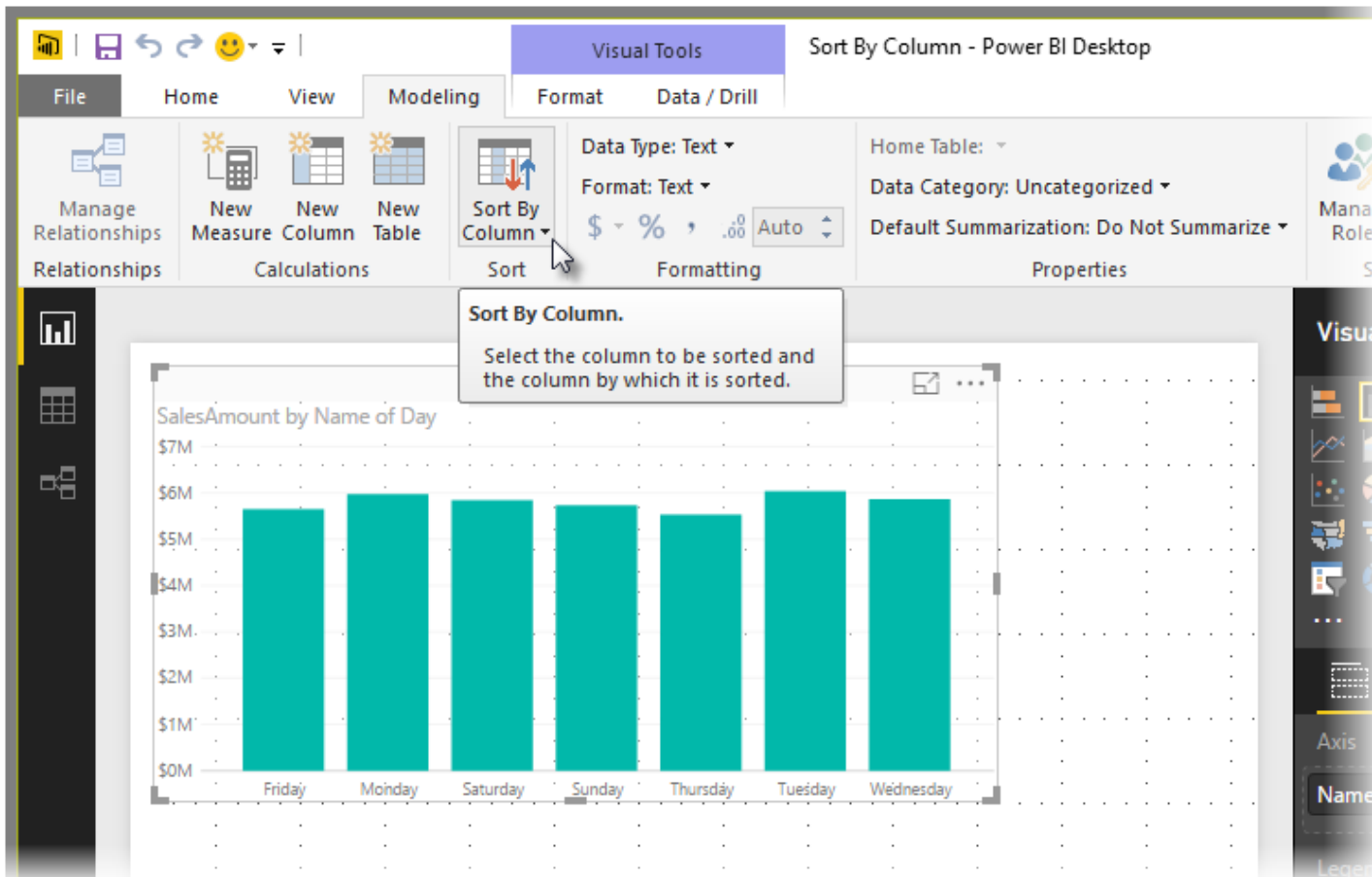


- Na imagem acima você notou a barra amarela ao lado de Custo de vida no menu Classificar Por, o que indica que a visualização estava classificada pela coluna Custo de vida.
- Classificar por outra coluna é fácil – basta selecionar as reticências para mostrar o menu Classificar Por e selecionar outra coluna.
- Ao selecionarmos a coluna Bem-estar da comunidade como sendo a coluna pela a qual queremos classificar a visualização é alterada ganhando uma nova forma.



Selecione a ordem de classificação - menor para o maior, maior para o menor

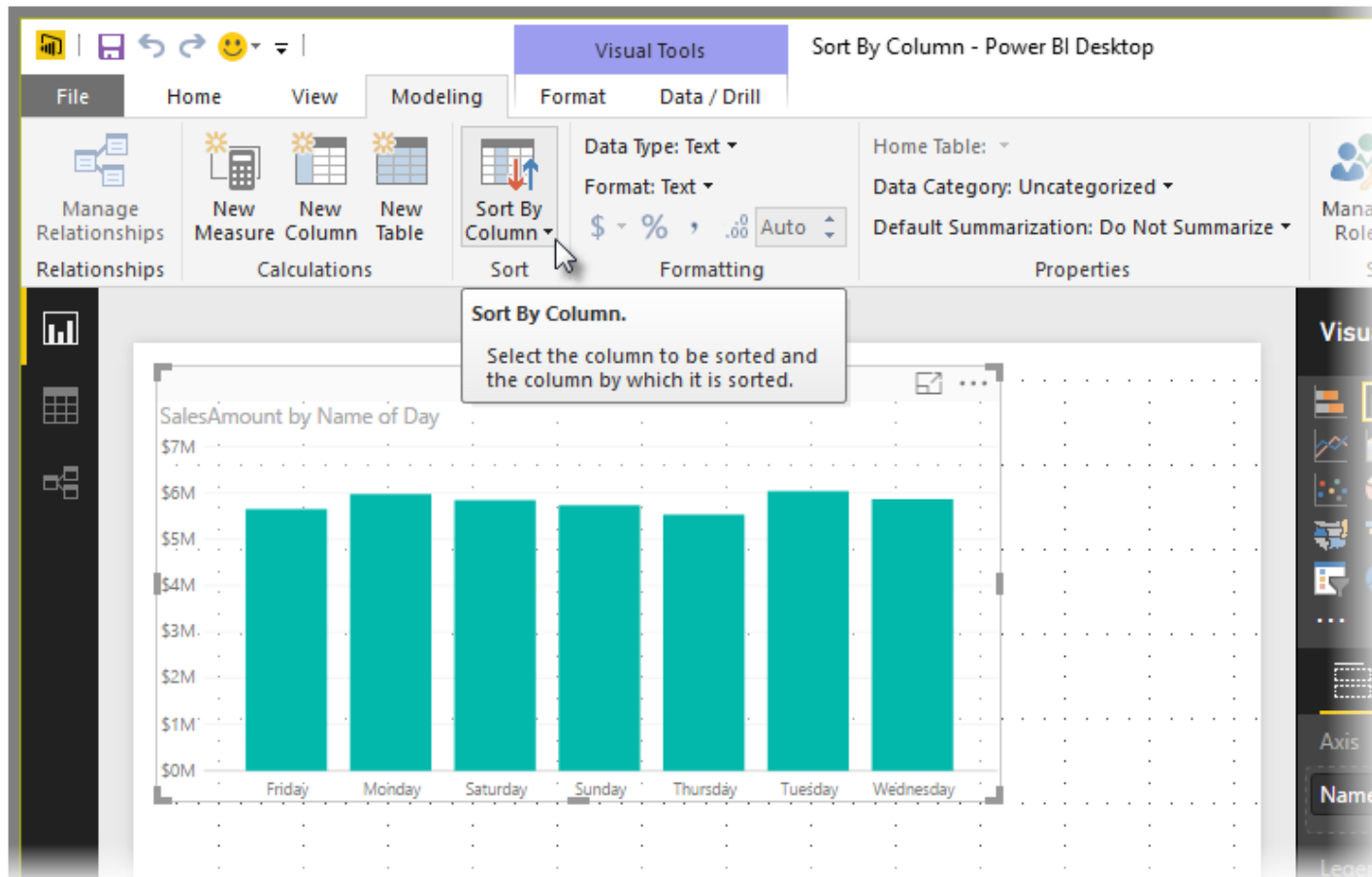
- De acordo com o site da Microsoft, quando analisamos de forma mais detalhada o menu Classificar Por da imagem anterior, podemos ver que o ícone ao lado de Classificar Por mostra Z/A (Z sobre A)
- Quando Z/A é exibido, significa que a visualização está sendo classificada pela coluna selecionada em ordem decrescente. Se quiser alterar basta tocar ou clicar no ícone Z/A e a ordem de classificação é alterada para A/Z, ordem crescente.



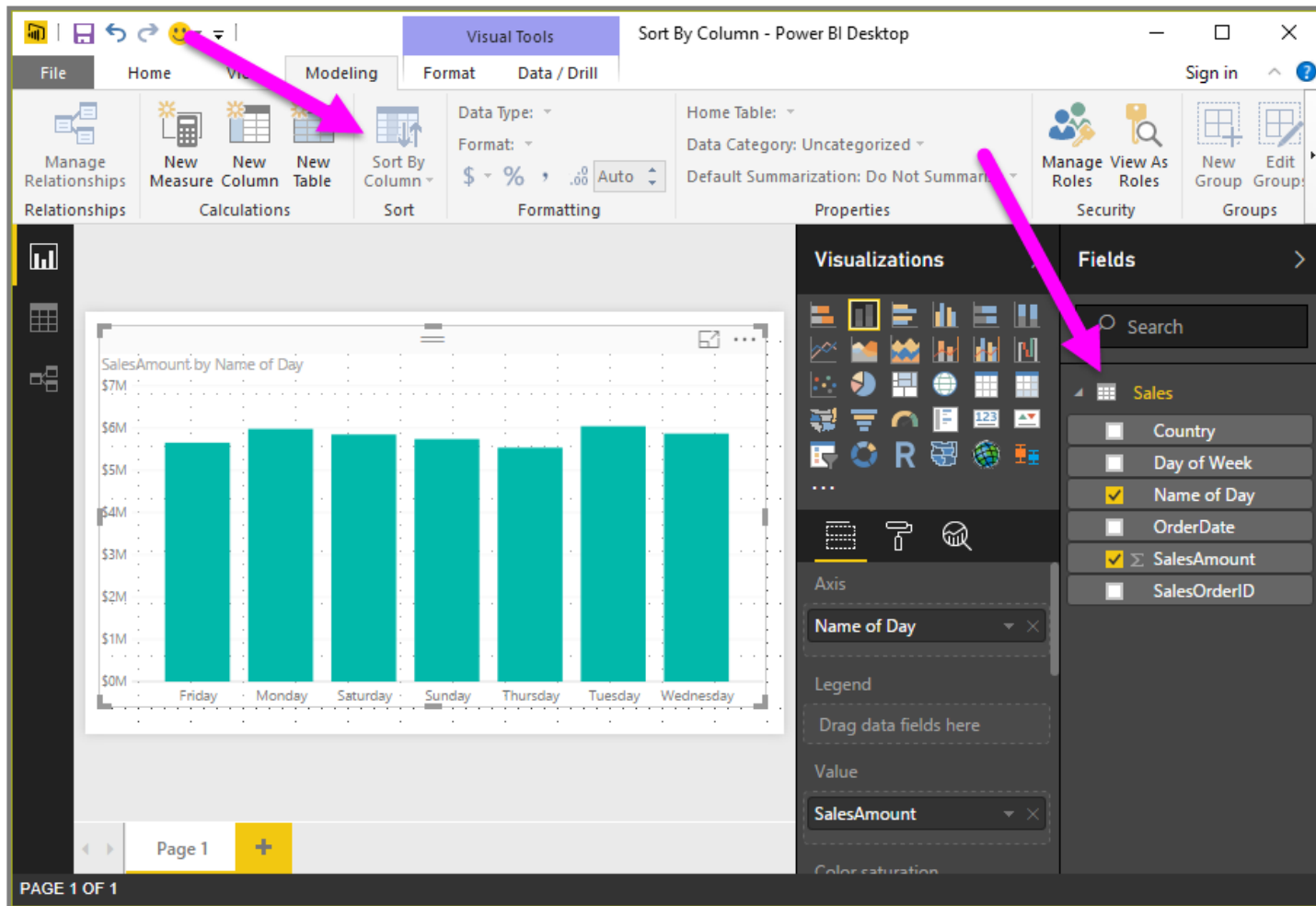
Classificar usando o botão Classificar por Coluna

- Existe outra maneira de classificar os dados, usando o botão Classificar por Coluna na área de opções Modelagem.

Classificar usando o botão Classificar por Coluna

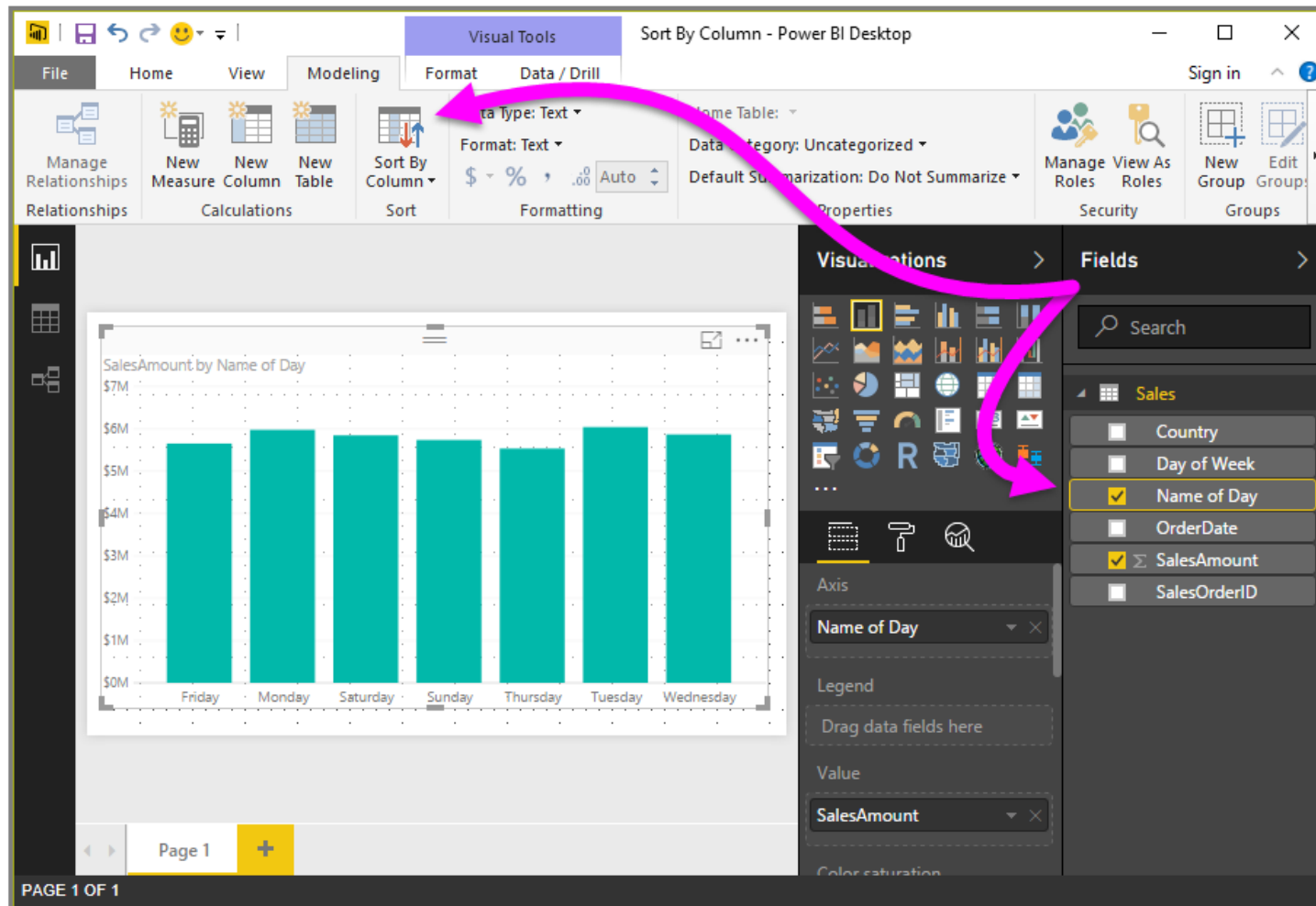


- Essa abordagem de classificação exige que você selecione uma coluna no painel Campos e, em seguida, selecione o botão Classificar por Coluna para escolher como (por qual coluna) você deseja classificar seu visual.
- Se a coluna não for selecionada o botão Classificar por Coluna fica inativo.
- Vejamos um exemplo comum:
 - você tem dados de cada dia da semana e deseja classificá-los com base em ordem cronológica.
- As etapas a seguir mostram como fazer isso.



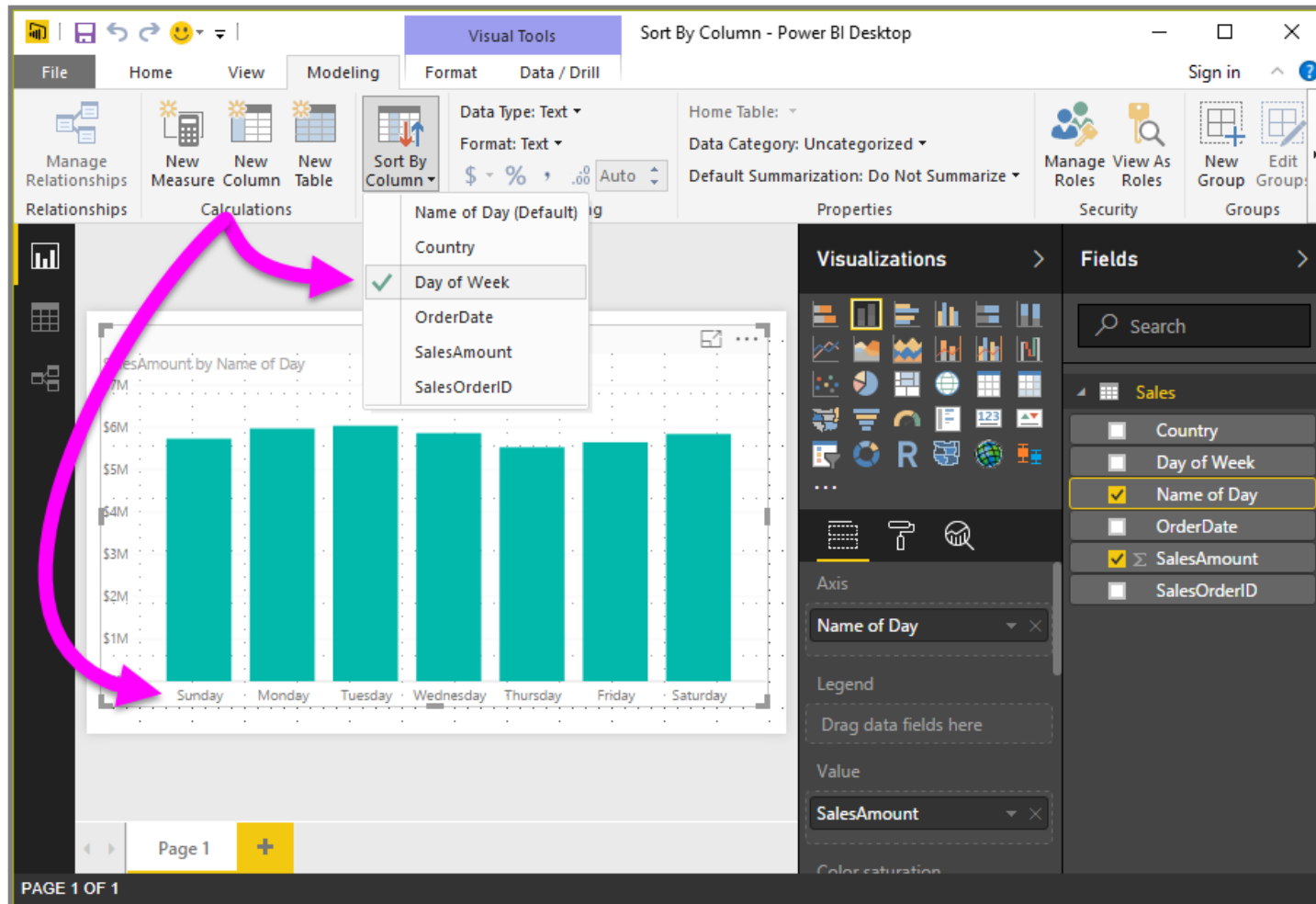
Classificar usando o botão Classificar por Coluna

- 1. Primeiro, observe que quando o visual é selecionado, mas nenhuma coluna é selecionada no painel Campos, o botão Classificar por Coluna fica inativo (esmaecido).



Classificar usando o botão Classificar por Coluna

- 2. Quando selecionamos a coluna pela qual queremos classificar no painel Campos, o botão Classificar por Coluna se torna ativo.



Classificar usando o botão Classificar por Coluna

- 3. Agora, com o visual selecionado, podemos selecionar Dia da Semana, em vez do padrão (Nome do Dia), e o visual classifica na ordem que queremos: por dia da semana.

ARREDONDAMENTO DE DADOS

Se o Algarismo a ser suprimido for:

- **Menor que 5: Basta suprimí-lo.**

Ex: 5,052 (Para um número centesimal) : 5,05

Ex: 103,701 (Para um número decimal):103,7

- **Maior que 5 ou igual a 5: Para suprimí-lo acrescente uma unidade ao algarismo que o precede.**

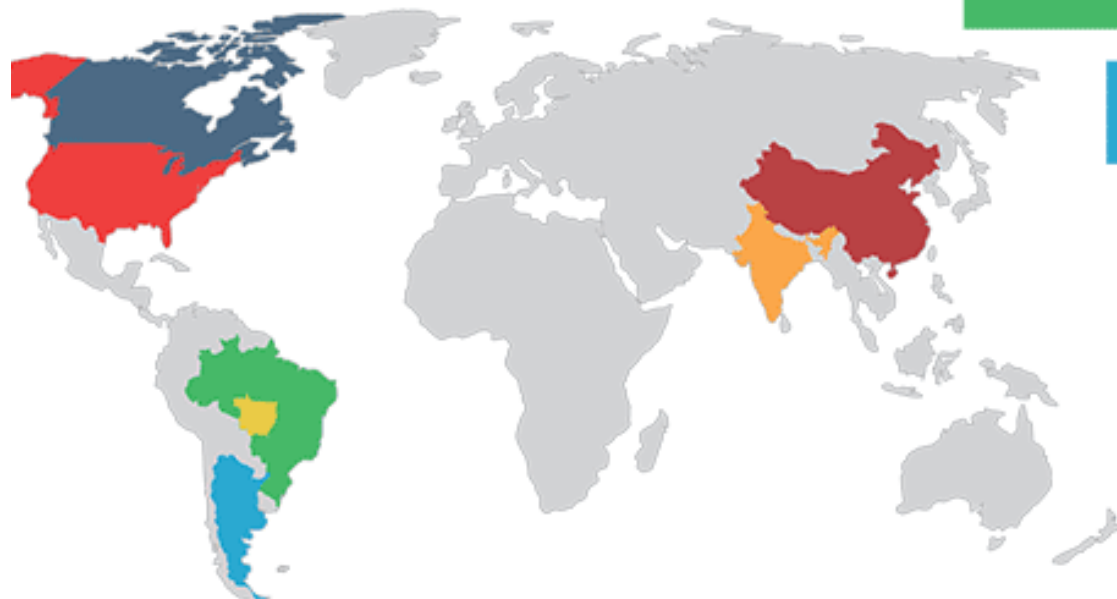
Ex: 5,057 (Para um número centesimal) : 5,06

Ex: 24,791 (Para um número decimal): 24,8

Arredondar dados no Power BI

- Antes de apresentar a funcionalidade de arredondamento vamos entender por que arredondamos dados
- Arredondamos dados para sermos mais ágeis na tomada de decisão e para não cometermos erros de cálculos.
- Números muito grandes ou então com muitas casas decimais tem grande chance de nos levar a cometer algum tipo de erro nos cálculos.
- Uma alternativa é utilizar o processo de arredondamento para deixar os números mais simples e fáceis de serem interpretados.

6: área plantada com transgênicos nundo (em milhões de hectares)



s plantadas:

ja, milho, algodão, canola,
a, papaia, alfaça, beterraba.
ioja, milho, algodão.
na: soja, milho, algodão.

Índia: algodão.

Mato Grosso: soja, milho, algodão.

Canadá: canola, milho, soja, beterraba.

China: algodão, papaia, alamo, tomate, pimentão.

so Brasil inclui os 11 milhões de ha do estado do Mato Grosso. Fonte: ISAAA 2014



Arredondar dados no Power BI

- Os números apresentados estão em milhões de hectares, conforme texto explicativo do título.
- Essa transformação facilitou a comparação e principalmente a visualização, imagine se em vez de arredondados os números fossem expressos em milhões, a visualização ficaria sobrecarregada.

ESTIMANDO POR ARREDONDAMENTO

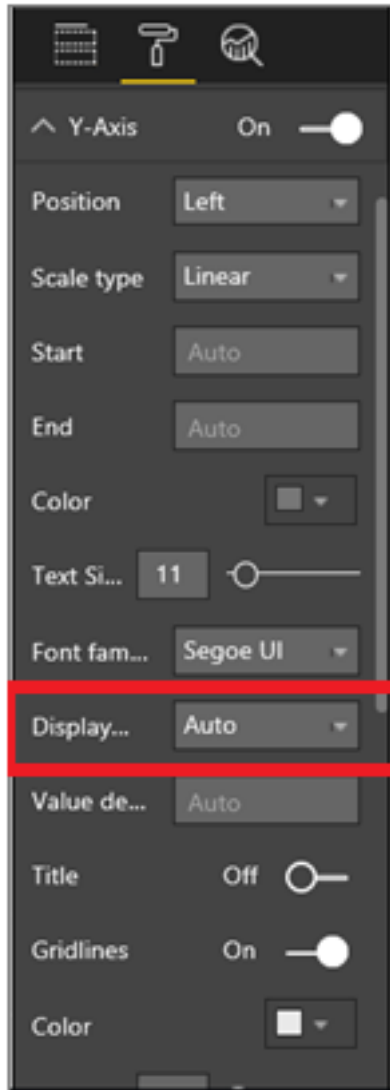
Você tem R\$ 100,00 para gastar em uma loja e quer saber rapidamente se o dinheiro é suficiente para comprar uma camiseta que custa R\$ 18,00, uma calça que custa R\$ 24,00 e um par de tênis de R\$ 47,00. Como fazer?

Arredondar os preços para a **dezena** mais próxima é uma alternativa:

18 para 20	}	$20 + 20 + 50 = 90$ (Valor exato 89)
24 para 20		
47 para 50		

Arredondar dados no Power BI

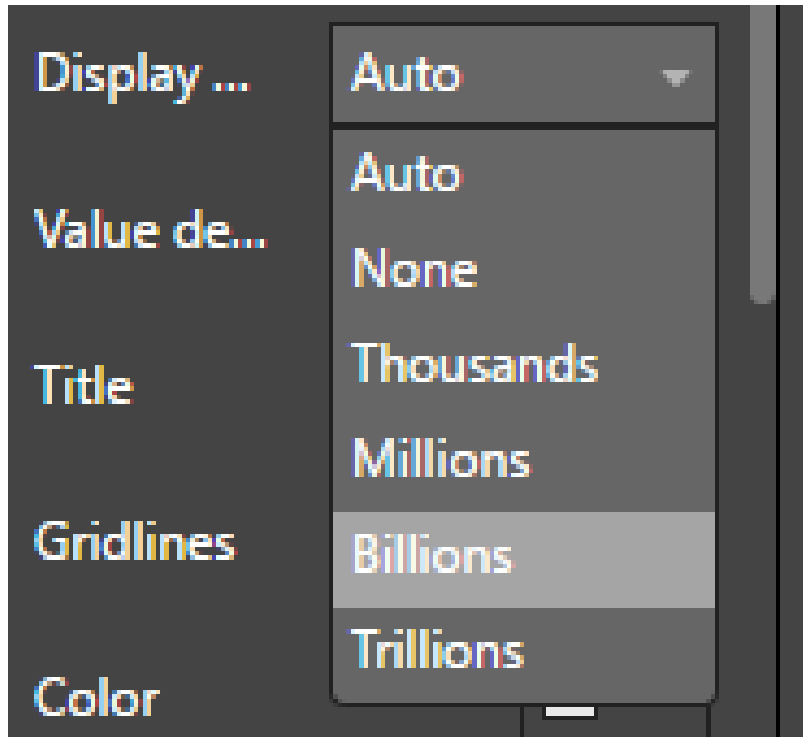
- Veja a imagem e confirme a facilidade de fazer cálculos com números arredondados



Arredondar dados no Power BI

- O ponto de atenção em arredondamentos são as perdas, números muito grandes podem apresentar perdas significativos, e números com muitas casas decimais podem perder precisão.
- No Power BI para você fazer o arredondamento é necessário entrar em customização do Eixo Y

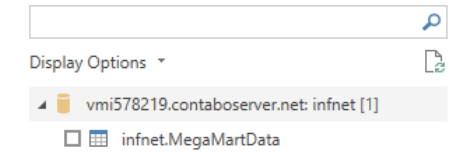
Arredondar dados no Power BI



- Selecione no “Display Unit” uma das opções de arredondamento. Segue lista das opções para esses dados
- Outra forma de fazer arredondamentos é através de colunas calculadas, essa modalidade nós veremos na etapa de manipulação de dados.

Praticando!

Navigator



Construir um dashboard utilizando boas práticas de visualização de dados, utilizando a base de dados “infnet” com os dados da tabela “MegaMartData” armazenados no MariaDB (vmi578219.contaboserver.net).

- Faça ajustes ou modificações na base de dados necessários para atingir seu objetivo.
- Se atente ao tipo da informação e como ela pode ser melhor representada no Power BI.
- Use figuras e imagens para ilustrar seu dashboard fazendo uso de áreas de textos.

- Totalização do montante de compras por cidade
 - Crie também uma visão hierarquizada para podermos ver o montante de vendas por cada loja
- KPI do ticket médio do montante de compras
 - Por cidade
 - Por gênero
- Somatório e média de vendas de todos os departamentos da rede varejista para o último ano disponível
- Média de vendas do departamento de eletrônicos e vestuário para o último ano disponível, mostrando uma linha de tendencia das vendas para os próximos 2 meses.
- Ranking dos tops 10 clientes que mais gastaram nas lojas.
- Ranking das top 10 lojas em vendas e suas localizações, ordem decrescente.

Na próxima aula...

Vamos continuar nosso estudo de visualização de dados no Power BI e veremos como definir e formatar escalas.

Apresentaremos como fazer análise exploratória dos dados utilizando as funções estatísticas do Power BI.