



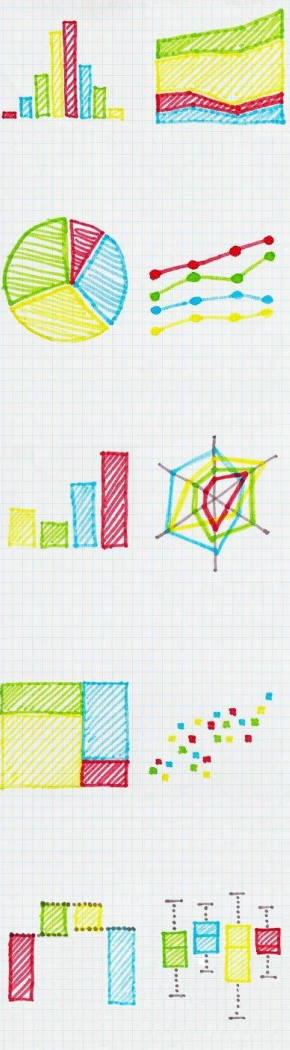
TIPOS DE GRÁFICOS PARA BUSINESS ANALYTICS





SUMÁRIO

INTRODUÇÃO	3
TIPOS DE VISUALIZAÇÃO	4
QUAL O MELHOR GRÁFICO?	20
CONCLUSÃO	21



Introdução

Perguntas surgem e você tem os dados. Criar um gráfico apresenta de forma simplificada a correlação entre os dois, mas muitas vezes você não tem certeza de que tipo de gráfico será melhor para visualizar as respostas.

Neste texto vamos ajudá-lo a escolher os melhores gráficos para o tipo de dados que você está analisando e as perguntas que você quer responder.

É importante saber que relacionar dados em gráficos estáticos isolados limita o número de perguntas que você pode responder.

Os seus dados devem ser a peça principal em uma tomada de decisão, usando-os para contar uma história. Combine gráficos relacionados. Adicione um mapa. Forneça filtros para uma análise mais detalhada.

Que benefícios consigo com isso? Informações úteis de negócios e respostas a perguntas com muito mais rapidez e eficiência.

Tipos de visualização

Diferentes tipos de visualização devem estar disponíveis em uma boa plataforma de Business Analytics e podem ser mostradas simultaneamente e serem atualizadas dinamicamente quando os filtros correspondentes são modificados.

As visualizações podem ser ativas ou inativas e feitas para refletir as muitas dimensões dos dados, permitindo o controle de atributos visuais tais como tamanho, cor, formato, etc.

[Tabela](#)

[Tabela cruzada](#)

[Tabela gráfica](#)

[Gráfico de barras](#)

[Gráfico de linhas](#)

[Gráfico de combinação](#)

[Gráfico de pizza](#)

[Gráfico de dispersão](#)

[Mapa gráfico](#)

[Mapa de árvore](#)

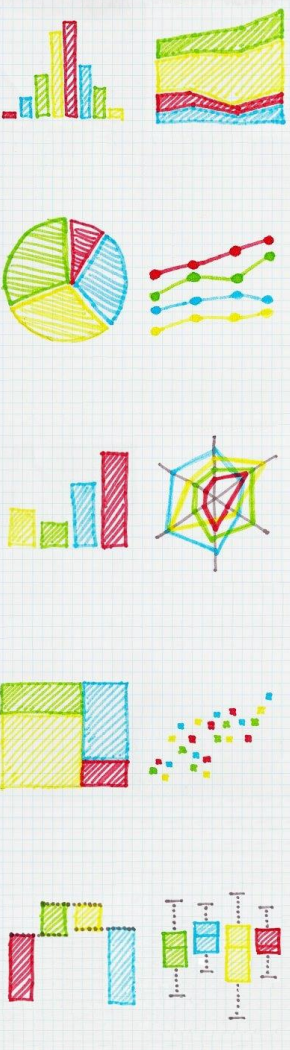
[Mapa de calor](#)

[Mapa de coordenadas paralelas](#)

[Tabela resumo](#)

[Visualização em grade](#)





Tabela

Uma tabela apresenta os dados em linhas e colunas e é usada para ver detalhes e comparar valores. É possível filtrar linhas não desejadas utilizando filtros e classificar a tabela, clicando no título de cada coluna.

É indicada para comparar dados entre categorias.

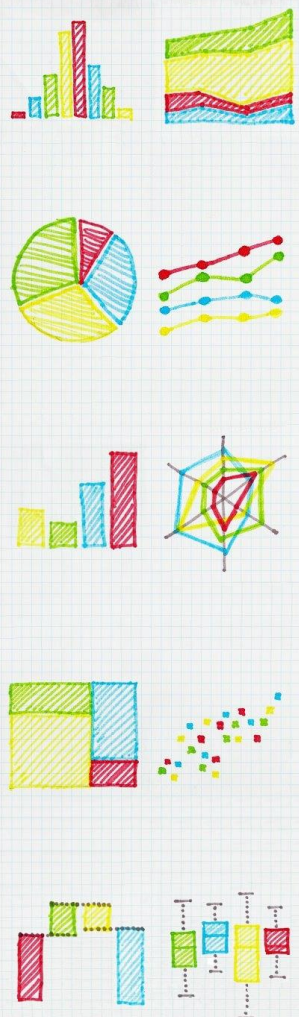
Continent	Country	City	Sales	Date of Sale
Asia	China	Beijing	1422	10/22/2002
Europe	Germany	Frankfurt	414	11/11/2002
Asia	China	Hong Kong	1556	11/25/2002
Asia	China	Beijing	6312	1/5/2003
Africa	Marocco	Casablanca	3556	5/17/2003
Africa	Marocco	Casablanca	6822	6/22/2003
Asia	Japan	Tokyo	288	11/1/2003
Europe	Germany	Berlin	2866	11/19/2003
Africa	Egypt	Alexandria	2673	12/24/2003
Africa	Marocco	Tangier	4231	2/29/2004
Asia	Japan	Yokohama	993	2/8/2004
Africa	Egypt	Cairo	3516	4/16/2004
Africa	Egypt	Alexandria	2323	5/5/2004

Tabela cruzada

Uma tabela cruzada apresenta dois caminhos de colunas e linhas. É também conhecida como tabela de pivô ou tabela multidimensional.

Sua melhor característica é poder estruturar, resumir e exibir uma grande quantidade de dados. Tabelas cruzadas também podem ser utilizadas para determinar se existe uma relação entre a linha variável e a coluna variável ou não. Como opção, a tabela cruzada pode exibir subtotais para colunas, e totais gerais para colunas, linhas ou para toda a medida.

Category	Type	2001	2002	2003	Grand total
Fruit	Apples	150	153	162	465
	Bananas	332	336	344	1012
	Pears	267	266	279	812
	Subtotal	749	755	785	2289
Vegetables	Cucumber	140	141	152	433
	Lettuce	246	245	258	749
	Tomatoes	156	161	168	485
	Subtotal	542	547	578	1667
Grand total		1291	1302	1363	3956



Os mesmos dados exibidos em uma tabela normal tomam muito mais espaço e é muito mais difícil de se obter uma visão geral:

Year	Month	Category	Type	Sales	Cost
2001	January	Fruit	Apples	12	10
2001	January	Fruit	Pears	21	13
2001	January	Fruit	Bananas	29	26
2001	January	Vegetables	Cucumber	9	6
2001	January	Vegetables	Tomatoes	13	11
2001	January	Vegetables	Lettuce	22	20
2001	February	Fruit	Apples	11	9
2001	February	Fruit	Pears	21	14
2001	February	Fruit	Bananas	31	27
2001	February	Vegetables	Cucumber	8	6
2001	February	Vegetables	Tomatoes	12	10
2001	February	Vegetables	Lettuce	20	19
2001	March	Fruit	Apples	9	10
2001	March	Fruit	Pears	19	13
2001	March	Fruit	Bananas	32	27
2001	March	Vegetables	Cucumber	8	6
2001	March	Vegetables	Tomatoes	11	11
2001	March	Vegetables	Lettuce	21	19
2001	April	Fruit	Apples	9	10
2001	April	Fruit	Pears	18	13
2001	April	Fruit	Bananas	32	27
2001	April	Vegetables	Cucumber	10	7
2001	April	Vegetables	Tomatoes	12	11
2001	April	Vegetables	Lettuce	21	19
2001	May	Fruit	Apples	10	9
2001	May	Fruit	Pears	20	13
2001	May	Fruit	Bananas	31	27
2001	May	Vegetables	Cucumber	11	7
2001	May	Vegetables	Tomatoes	13	11
2001	May	Vegetables	Lettuce	21	19
2001	June	Fruit	Apples	9	9

Tabela gráfica

Uma tabela gráfica é uma visualização resumida, projetada para fornecer uma grande quantidade de informações.

Pode ser configurada para exibir colunas com itens dinâmicos tais como *sparklines**, valores calculados ou ícones condicionais. Um valor é exibido para cada linha como especificado nos eixos das linhas. Também pode conter controles interativos que afetam outras partes da análise ou exibir outra página da análise.

Stock	Sparkline	Average Price	Cheapest
Barkley Books Ltd.		139.04	
Cameo Cars Inc.		164.58	
Flinch Foods Inc.		103.96	☆
Machton Machinery Ltd.		213.54	

(*) *Sparklines* são gráficos de linha simples e pequenos, tradicionalmente usados para exibir tendências ou variações de algumas variáveis:



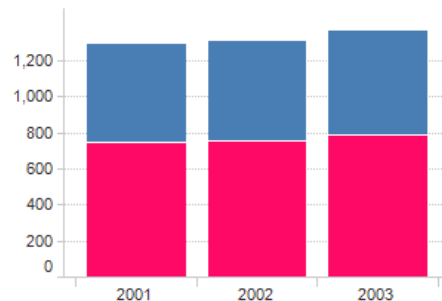
Gráfico de barras

Um gráfico de barras é uma forma de resumir um conjunto de dados categóricos. Ele mostra os dados, utilizando um número de barras de mesma largura, cada uma delas representando uma categoria particular. A altura de cada barra é proporcional a uma agregação específica (ex.: a soma dos valores na categoria que representa). As categorias podem ser algo como um grupo de idade ou uma localização geográfica.

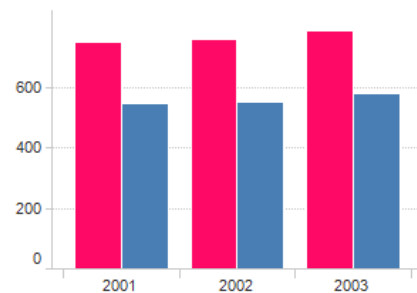
Se aplicado quando uma análise for criada, o gráfico de barras pode exibir informação adicional em linhas de referência ou em diferentes tipos de curvas. Essas linhas ou curvas podem, por exemplo, exibir quão bem os seus dados se adaptam a certo ajuste de curva polinomial ou para resumir uma coleção de pontos de dados amostrais ajustando-os a um modelo que descreverá os dados e exibirá uma curva ou uma linha reta no topo da visualização. A curva normalmente modifica a aparência dependendo de quais valores você filtrou na análise.

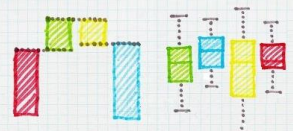
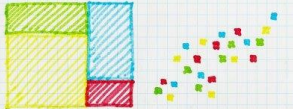
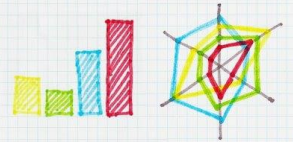
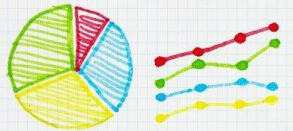
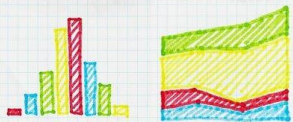
Exemplo: Uma tabela de dados contém os números de venda para um número de diferentes frutas e legumes. O gráfico de barras pode mostrar a soma total das vendas de diferentes anos.

As barras são coloridas pela Categoria coluna a qual contém os dois valores de Fruta e Vegetais.

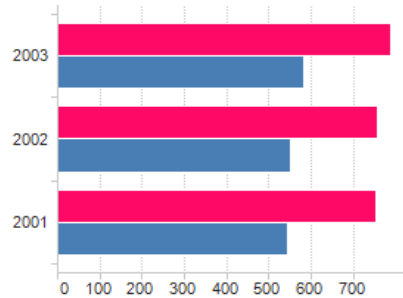


Na segunda figura, a coluna de Categoria foi adicionada ao eixo X, tal que as barras foram divididas em duas barras separadas, uma para cada categoria.





As barras podem também ser mostradas horizontalmente:



Barras de erro

As barras de erro são utilizadas para indicar o erro estimado em uma medição. Em outras palavras, uma barra de erro indica a incerteza de um valor. Utilizada em um gráfico de barras, as barras de erro se parecem com isso:

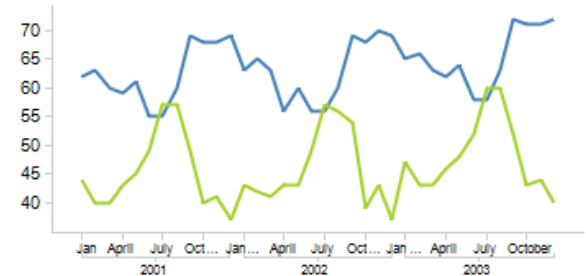


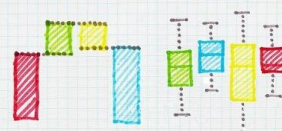
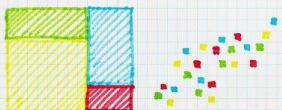
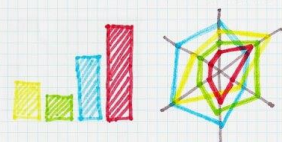
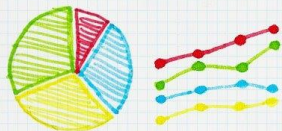
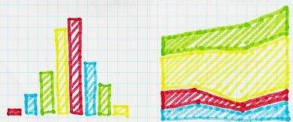
Gráfico de linhas

Gráficos de linhas são ideais para exibir tendências ao longo do tempo.

Um exemplo padrão seria como as ações de uma certa empresa se comportam ao longo do tempo no mercado da bolsa. Entretanto, não necessariamente precisa ser o tempo ao longo do eixo X. Qualquer dado que se comporte como uma função com respeito a variável no eixo X pode ser plotado.

Gráficos de linha enfatizam o fluxo de tempo e taxa de mudança mais do que a quantidade de mudanças. Múltiplas escalas podem ser utilizadas no eixo Y para quando você quiser comparar diversas linhas com variações de valores significativamente diferentes.





Se aplicado quando uma análise é criada, o gráfico de linha pode exibir informação adicional em linhas de referência ou em diferentes tipos de curvas. Essas linhas ou curvas podem, por exemplo, exibir quão bem os seus dados se adaptam a certo ajuste de curva polinomial ou para resumir uma coleção de pontos de dados amostrais ajustando-os a um modelo que descreverá os dados e exibirá uma curva ou uma linha reta no topo da visualização.

Barras de erro

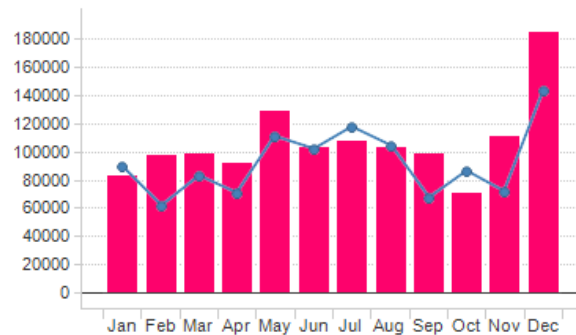
As barras de erro são utilizadas para indicar o erro estimado em uma medição. Em outras palavras, uma barra de erro indica a incerteza de um valor. Utilizada em um gráfico de linha, as barras de erro se parecem com isso:



Gráfico de Combinação

O gráfico de combinação é uma visualização que combina as características do gráfico de barras com as do gráfico de linha.

O gráfico de combinação mostra os dados utilizando um número de barras e linhas, cada qual representando uma categoria particular. Uma combinação de barras e linhas na mesma visualização pode ser útil quando comparar valores em diferentes categorias, desde que a combinação forneça uma visão clara de qual categoria está acima ou abaixo. Um exemplo disso pode ser visto ao utilizar o gráfico de combinação para comparar as vendas projetadas com as vendas atuais para diferentes períodos de tempo.



O exemplo acima contém duas séries: as barras representam as vendas projetadas para cada mês e a linha representa as vendas atuais para os mesmos meses.

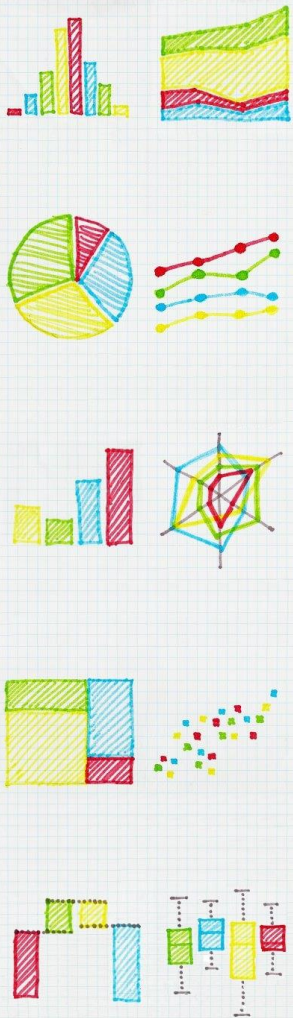


Gráfico de pizza

Gráficos de pizza são elementos circulares divididos em setores. Cada setor da pizza exibe o tamanho de uma parte da informação relacionada. Gráficos de pizza normalmente são utilizados para exibir os tamanhos relativos das partes de um todo.

Countries by Area

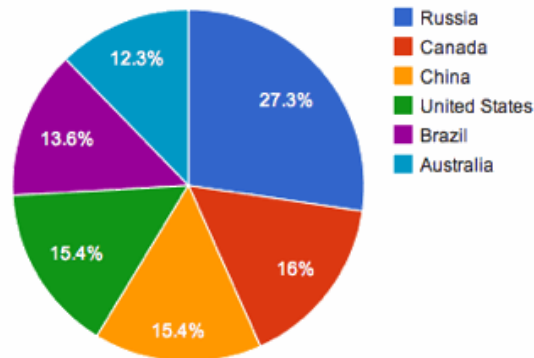
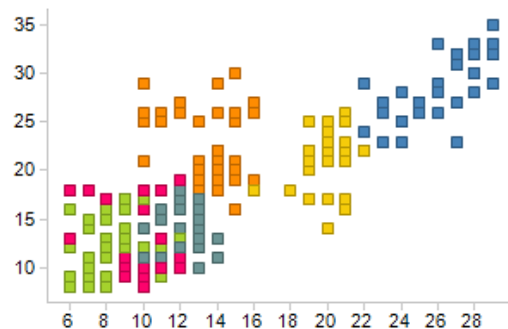
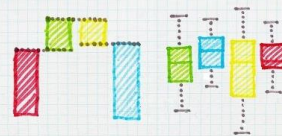
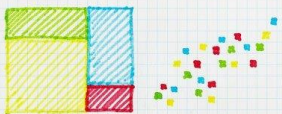
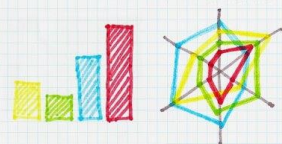
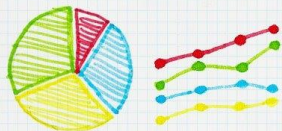
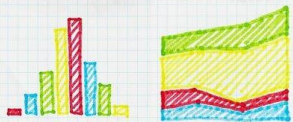


Gráfico de dispersão

Gráficos de dispersão são utilizados para pontuar dados em um eixo vertical e horizontal com a intenção de exibir quanto uma variável é afetada por outra.

Cada linha na tabela de dados é representada por um marcador cuja posição depende dos seus valores nas colunas determinados nos eixos X e Y. Múltiplas escalas podem ser utilizadas no eixo Y para quando você quiser comparar diversos marcadores com faixas de valores significativamente diferentes. Uma terceira variável pode ser configurada para corresponder a uma cor ou ao tamanho (por ex., um gráfico de bolhas) dos marcadores, e assim adicionar outra dimensão ao gráfico.





No gráfico de dispersão acima, promoções (Y) está plotado em função de custo (X) para um número de produtos diferentes (coloridos por produto), para exibir uma baixa correlação positiva.

A relação entre duas variáveis é chamada de correlação. Se os marcadores estão próximos a formar uma linha reta no gráfico de dispersão, as duas variáveis possuem uma alta correlação. Se os marcadores estiverem igualmente distribuídos no gráfico de dispersão, a correlação é baixa, ou zero.

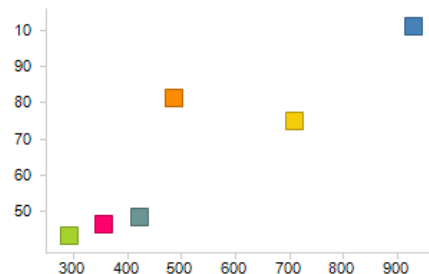
Entretanto, mesmo se a correlação pareça estar presente, esse pode não ser o caso. Ambas as variáveis podem estar relacionadas a uma terceira variável, então expandir a sua variação ou uma pura coincidência pode causar uma aparente correlação.

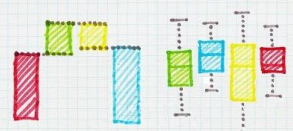
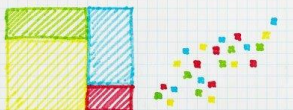
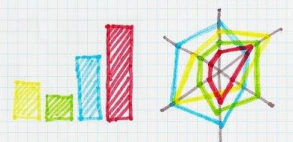
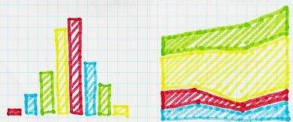
Se aplicado quando uma análise é criada, o gráfico de dispersão pode exibir informação adicional em linhas de referência ou em diferentes tipos de curvas. Estas linhas ou curvas podem, por exemplo, exibir quão bem os seus dados se adaptam a certo ajuste de curva polinomial ou para resumir uma coleção de pontos de dados amostrais ajustando-os a um modelo que

descreverá os dados e exibirá uma curva ou uma linha reta no topo da visualização. A curva normalmente modifica a aparência dependendo de quais valores você filtrou na análise.

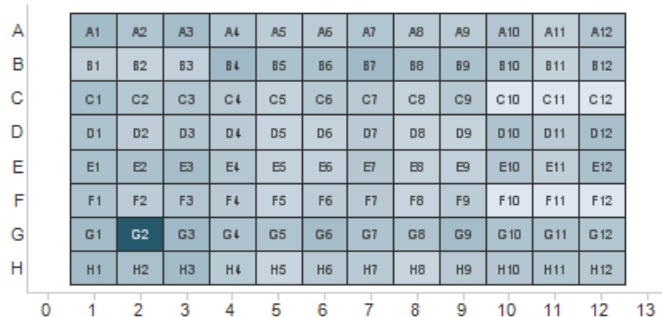
O gráfico de dispersão pode também ser usado juntamente com agregação, por exemplo: Soma ou Média. Neste caso, os valores para uma certa categoria são agrupados, para exibir um único marcador para cada categoria. Os marcadores agregados podem também ser dimensionados pela contagem dos itens dentro de cada categoria ou por qualquer outra coluna.

Os marcadores não mostram a Soma das Vendas para cada produto.





Outra opção é utilizar marcadores lado a lado. Isso significa que todos os marcadores terão o mesmo tamanho e serão exibidos em um layout similar ao de grade como visto no exemplo abaixo:



Esse exemplo exibe os resultados de um experimento conduzido em uma placa de ensaio onde cada marcador no gráfico de dispersão representa um poço no ensaio de placa e os marcadores coloridos representam os resultados do experimento para cada um dos poços.

Barras de erro

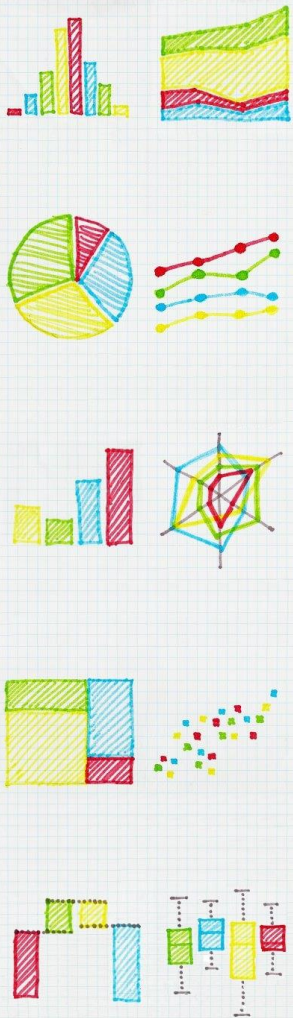
As barras de erro são utilizadas para indicar o erro estimado em uma medição. Em outras palavras, uma barra de erro indica a incerteza de um valor. Utilizada em um gráfico de dispersão, as barras de erro se parecem com isso:



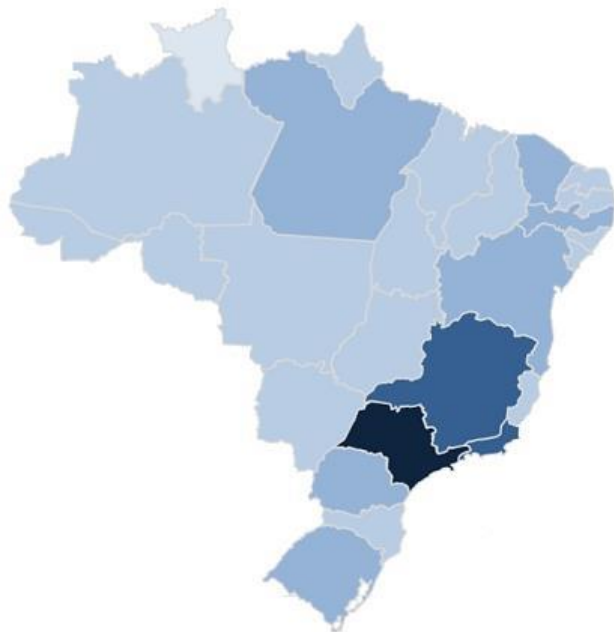
Mapa gráfico

Um gráfico de mapa é usado para mostrar itens em um fundo, que é quase sempre geográfico.

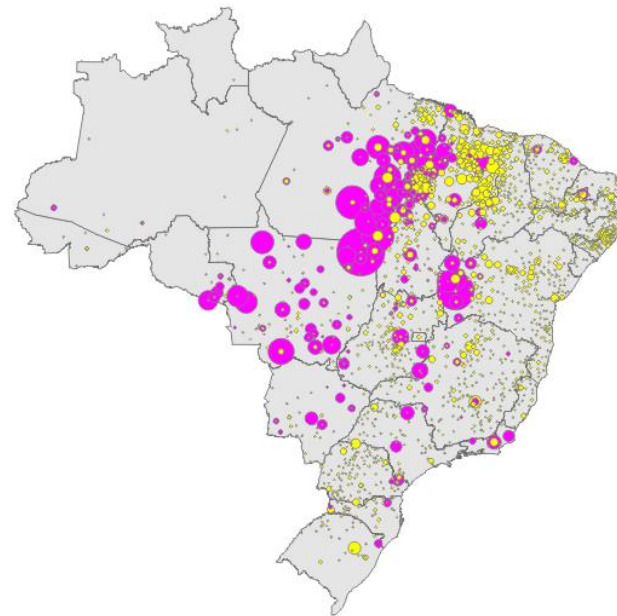
Os mapas gráficos podem conter diferentes tipos de informação em uma ou mais camadas diferentes. Podem conter formas interativas e exibir marcadores de diferentes tipos em uma imagem ou no fundo do mapa.

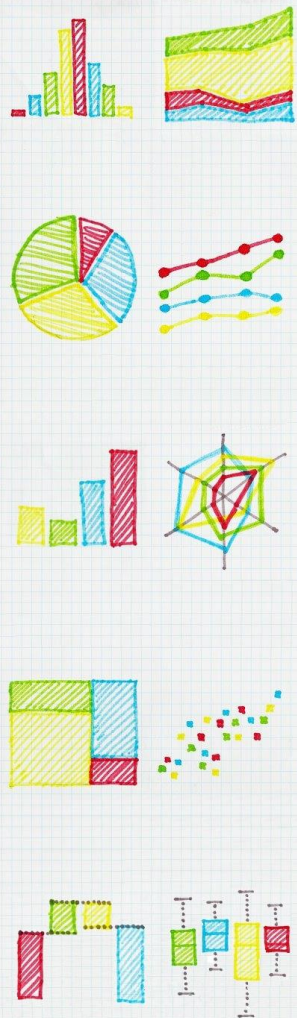


Abaixo está um exemplo do mapa gráfico com uma camada que contém formas interativas. As diferentes partes do mapa, neste caso o Brasil, podem ser interagidas da mesma forma que você faria com itens em qualquer outra visualização.

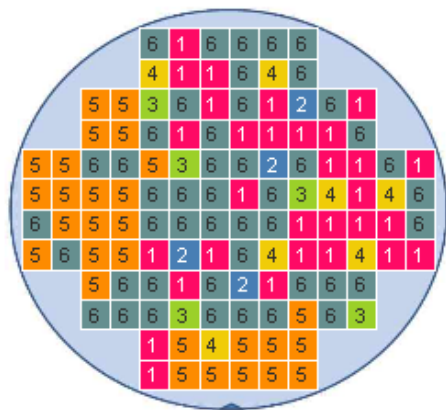


No exemplo abaixo, um mapa gráfico com marcadores é usado para exibir os marcadores em um contexto geográfico ou espacial.





Um mapa gráfico também pode ser utilizado para exibir dados que não são geográficos. O exemplo abaixo mostra diferentes tipos de falhas em um *wafer*, um material semiconductor utilizado para manufaturar microchips.



Mapa de árvore

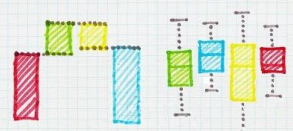
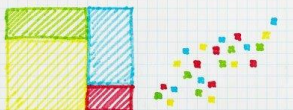
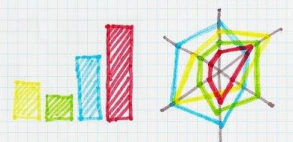
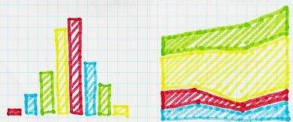
Mapas de árvores são ideais para exibir grandes quantidades de dados hierarquicamente estruturados (estruturados em árvore).

O espaço na visualização é dividido em retângulos que são dimensionados e ordenados por uma variável quantitativa. Os níveis na hierarquia do mapa de árvore são visualizados como retângulos contendo outros retângulos.

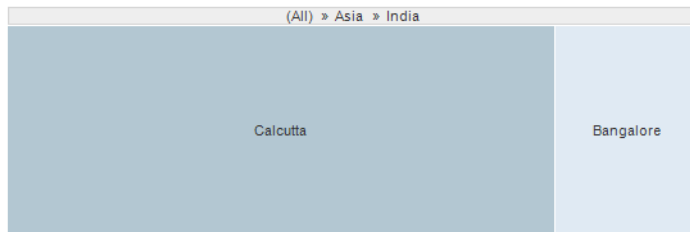
Cada conjunto de retângulos no mesmo nível na hierarquia representa uma coluna ou uma expressão na tabela de dados. Cada retângulo individual em um nível na hierarquia representa uma categoria em uma coluna.

Exemplo: um retângulo representando um continente pode conter diversos retângulos representando países neste continente. Cada retângulo representando um país pode, por sua vez, conter diversos retângulos representando cidades nestes países.

Africa		Europe		Asia
Marocco	South Africa	France	Germany	China
Casablanca	Johannesburg Cape Town	Cannes	Frankfurt	Beijing
	Egypt			Hong Kong
Tangier	Alexandria Cairo	Paris	Berlin	India
				Calcutta
				Bangalor



Para tomar um olhar mais próximo de uma determinada parte do mapa de árvore, é possível navegar de um nível de hierarquia mais alto para um mais baixo.



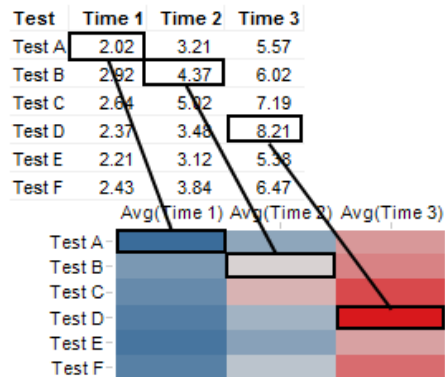
Mapa de calor

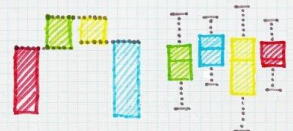
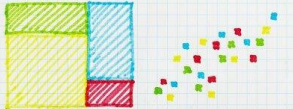
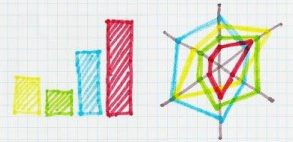
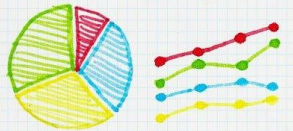
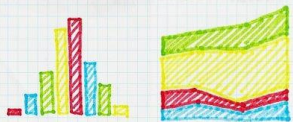
A melhor forma de entender um mapa de calor é pensar em uma tabela ou planilha que contém cores ao invés de números.

O gradiente de cor padrão configura o menor valor no mapa de calor como azul escuro, o maior valor como um vermelho brilhante e valores medianos como cinza claro, com uma transição correspondente (ou gradiente) entre estes extremos.

Os mapas de calor são bastante apropriados para visualizar grandes quantidades de dados multidimensionais e podem ser utilizados para identificar grupos de linhas com valores similares, conforme elas são mostradas nas áreas de cor similar.

O exemplo abaixo ilustra como os valores na tabela estão dispostos como um gradiente de cor nas células do mapa de calor.





Dendrograma

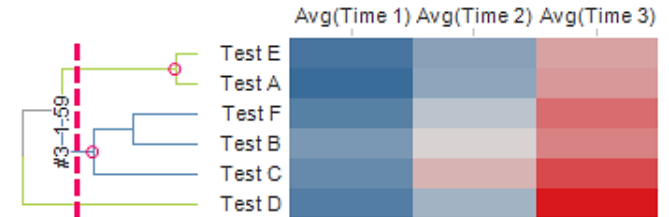
Às vezes é útil combinar os mapas de calor com um agrupamento hierárquico, que é uma maneira de arranjar itens em uma hierarquia baseada na distância ou similaridade entre eles.

O resultado de um cálculo de agrupamento hierárquico está exibido em um mapa de calor como um dendrograma, que é uma estrutura em árvore da hierarquia. Dendrogramas de linhas mostram a distância (ou similaridade) entre as linhas e cujos nós se apresentam como resultado de um cálculo de agrupamento.

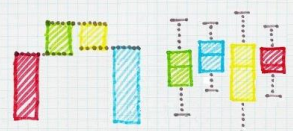
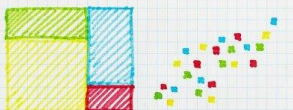
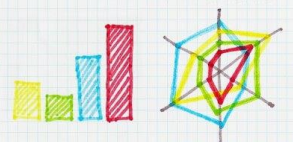
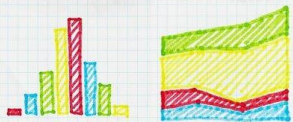
Dendrogramas de coluna mostram a distância (ou similaridade) entre as variáveis (as colunas selecionadas de valores em células).

O dendrograma torna fácil destacar e marcar algo no mapa de calor. É possível destacar grupos e suas células correspondentes no mapa de calor. Você pode clicar para marcar um grupo. Isso irá marcar as células correspondentes no mapa de calor.

O exemplo abaixo mostra um mapa de calor com um dendrograma de linha onde as distâncias entre as linhas foram calculadas. A linha pontilhada vertical é a linha de poda, a qual pode ser arrastada lateralmente no dendrograma.



Como um resultado do cálculo de agrupamento, as linhas no mapa de calor foram reordenadas para corresponder ao cálculo do grupo. Teste A e Teste E estão posicionados no mesmo grupo. Teste F e Teste B estão posicionados juntos em um outro grupo e este grupo forma outro grupo juntamente com o Teste C.



O Teste D não está incluído em nenhum destes grupos. Isso indica que Teste A e Teste E estão mais próximos um do outro do que eles estão do Teste F, Teste B, Teste C ou Teste D. Isso também indica que Teste D é o que está mais distante das outras linhas.

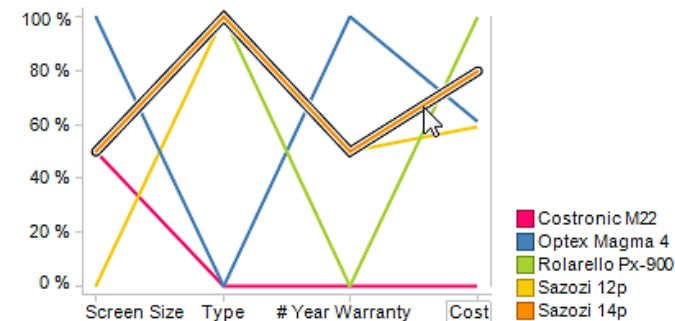
Gráfico de coordenadas paralelas

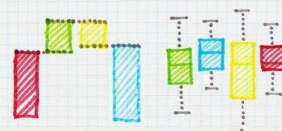
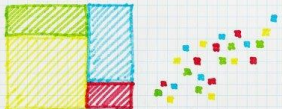
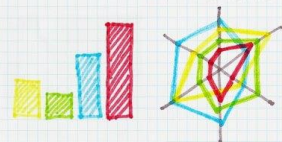
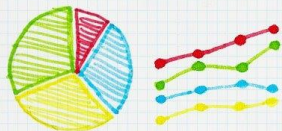
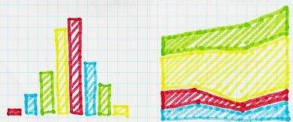
Em um gráfico de coordenadas paralelas, cada fileira em uma tabela de dados é uma linha ou perfil. Cada atributo de uma fileira é representada por um ponto na linha. Isso marca gráficos de coordenadas paralelas similares em aparência com gráficos de linhas, mas a forma como os dados são traduzidos em um gráfico é substancialmente diferente.

Considere, por exemplo, uma tabela de dados contendo informação sobre modelos de monitores de computador. Os dados nas várias colunas são de tipos completamente diferentes e totalmente incomparáveis uns com o outros.

Monitor Model	Screen Size	Type	# Year Warranty	Cost
Costronic M22	19"	CRT	1	\$150.00
Sazozzi 12p	17"	TFT	2	\$295.00
Sazozzi 14p	19"	TFT	2	\$345.00
Optex Magma 4	20"	CRT	3	\$299.00

Entretanto, os dados fornecem perfis para cada monitor. Estes perfis podem ser úteis para encontrar um monitor que corresponda às suas necessidades. O Gráfico de coordenadas paralelas é bastante útil para visualizar esses perfis.





Os valores em um gráfico de coordenadas paralelas são sempre normalizados e o valor mais baixo está plotado como 0% e o maior como 100%. Isso significa que para cada ponto ao longo do eixo X, o menor valor na coluna correspondente está configurado para 0% e o maior valor naquela coluna está configurado para 100% ao longo do eixo Y. A escala de várias colunas é totalmente separada, então não compare a altura da curva em uma coluna com a altura da curva em outra coluna.

O monitor Sazozzi 14p (o perfil laranja com o cursor em cima dele) é uma tela plana TFT de um tamanho médio (nesta seleção). Ela possui um período de garantia médio e seu preço está no campo superior. Esta pode ser a melhor compra, dependendo do caso de uso.

Tabela resumo

A tabela resumo é uma visualização que exibe numericamente uma informação estatística.

A informação é baseada em uma tabela de dados.

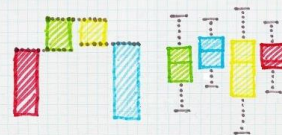
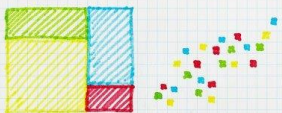
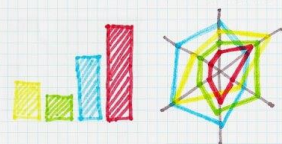
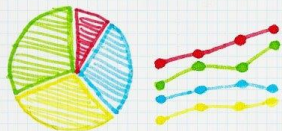
Quando você altera o conjunto de linhas filtradas, a Tabela Resumo atualiza automaticamente os valores exibidos para refletir a seleção atual.

Column	Sum	Avg	Min	Max	Median	StdDev
Sales	3956	18	8	35	18	7
Cost	3194	15	6	29	13	6

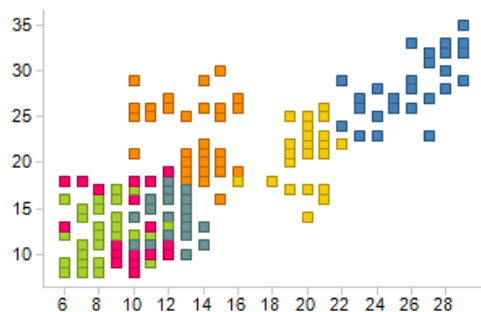
Visualização em grade

Visualizações em grade habilitam você a rapidamente reconhecer similaridades ou diferenças entre diferentes categorias nos dados.

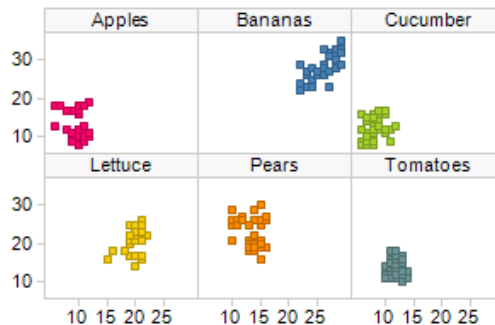
Cada painel individual em uma visualização em grade exibe um subconjunto de uma tabela original, onde os subconjuntos são definidos pelas categorias disponíveis em uma coluna ou hierarquia.

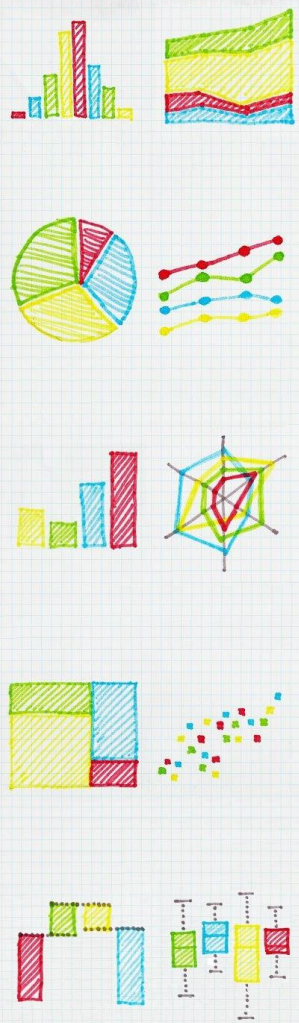


No gráfico de dispersão abaixo, promoções (Y) está plotado em função de custo (X) para um número de produtos diferentes (coloridos por produto), para exibir uma baixa correlação positiva.

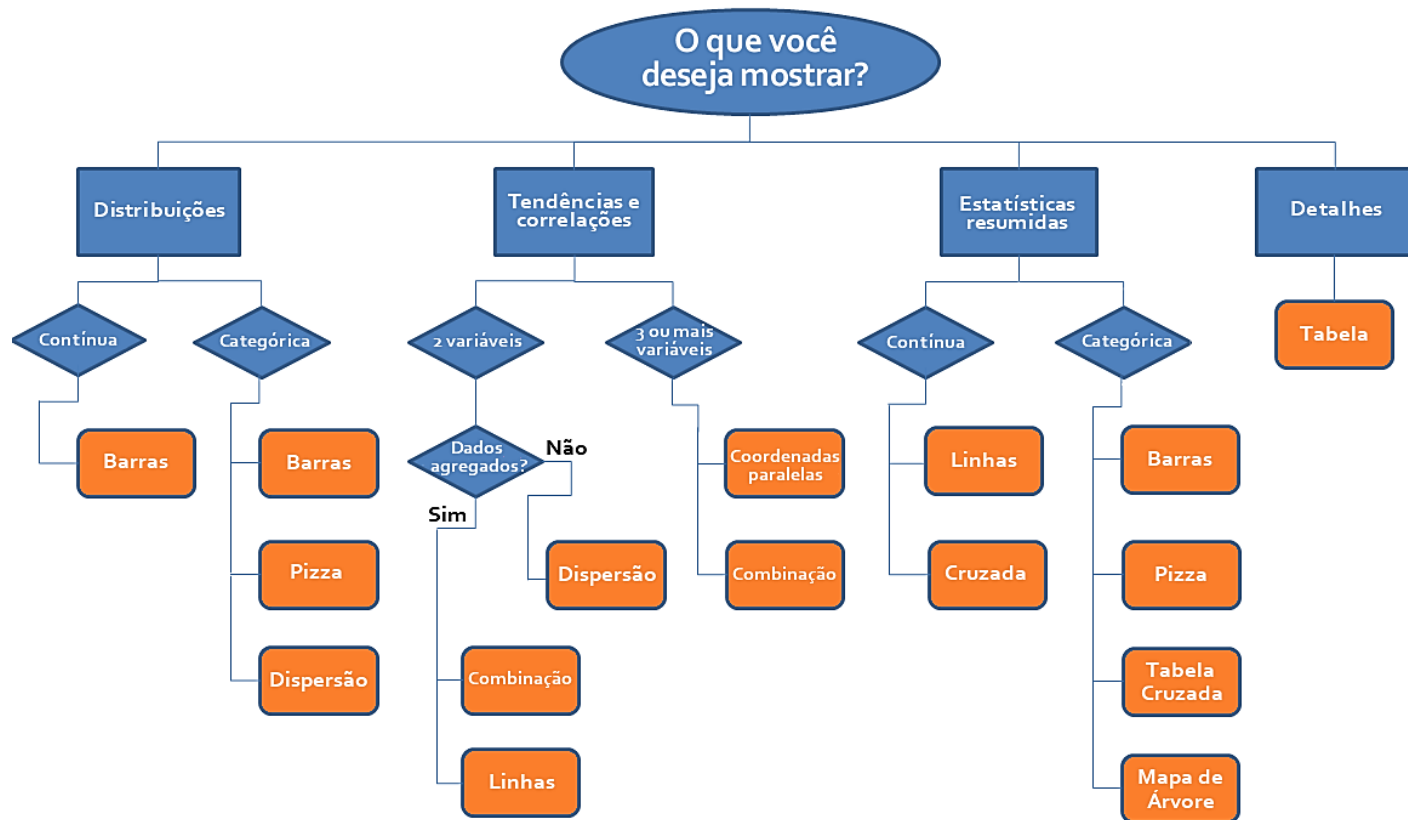


Cada produto pode ser exibido separadamente utilizando o treliçamento.





Qual o melhor gráfico?



Conclusão

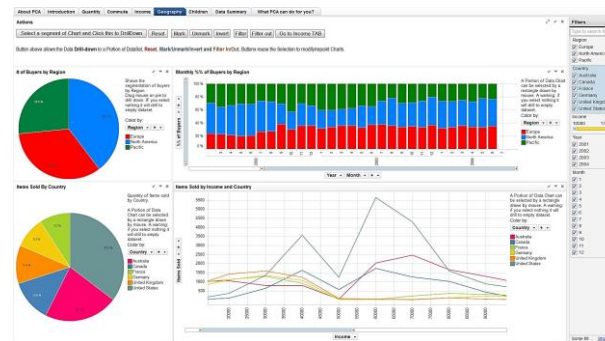
É importante estabelecer uma conexão entre os dados que você tem e as perguntas que você quer responder.

Selecionar o tipo certo de gráfico é a primeira etapa para obter as informações úteis que seus dados podem fornecer, mas eles podem ir muito além.

Seus gráficos devem ter interpretação intuitiva e oferecer interatividade em tempo real pois, à medida que você vê as aplicações de seus dados, pode continuar analisando as informações para fazer mais perguntas, que juntamente com fluxos de trabalho guiados irão ajudar você a responder a eventos que afetam seu negócio.

Os *dashboards* e visualizações de dados devem ser uma combinação de utilidade, usabilidade e beleza, a fim de impulsionar o engajamento e a ação de todos os envolvidos na organização.

Combinar gráficos em um painel (*dashboard*) é a melhor forma de criar um ambiente produtivo em que você pode não apenas perguntar, mas responder, compartilhar e distribuir informação dentro e fora da organização, em tempo real.





SOMOS UMA EMPRESA COM MAIS DE 25 ANOS, ESPECIALIZADA EM **BPM (GESTÃO DE PROCESSOS DE NEGÓCIO)**, **INTEGRAÇÃO DE SISTEMAS E BUSINESS ANALYTICS**.

☎ (11) 3044-2922

✉ roberto@officeware.com.br

🏠 officeware.com.br