

## O template de dispersão

Neste vídeo, vou falar mais a fundo sobre o template de dispersão, e este é um template que é possivelmente o mais amplo em seus usos do que qualquer template do Flourish. Nós sempre brincamos quando estamos desenvolvendo um novo template que "ah, talvez possamos apenas adicionar isso ao gráfico de dispersão", porque ele pode fazer tantas coisas.

E eu queria começar aqui no seletor de template, porque eu queria mostrar visualmente quantas coisas diferentes o template de dispersão pode fazer. Então, esses são todos os pontos de partida que temos com dispersão, e estes são todos os diferentes tipos de gráficos, geralmente de fato diferentes tipos de gráficos que você pode fazer com o template de dispersão. Acho que o único que não está aqui que você pode fazer é o diagrama "beeswarm". Então vou passar por eles e explicar a estrutura básica do template de dispersão e depois mostrar como cada um dos outros diagramas são feitos. Mas saiba que se você quiser fazer um destes, você não precisa fazê-lo 100%, apenas jogando seus dados neste gráfico básico, depois no gráfico de dispersão, e depois tentar fazer as configurações você mesma nas configurações de coluna. Existem esses pontos de partida que devem ser bons guias.

Então eu vou começar aqui em um gráfico de dispersão básico que está mostrando, eu acho, o PIB per capita e a expectativa de vida de vários países. Cada um desses pontos é um país. E eu queria mostrar isso porque eu quero mostrar as diferenças entre um gráfico de dispersão básico e um gráfico de bolhas. Aqui está o gráfico de dispersão básico, também é bom saber que no template de dispersão, o painel de configurações é meio enorme e um pouco intimidador. Mas não se assuste muito com isso. Muitas dessas coisas, se você não está fazendo esse tipo de gráfico, você não precisa mexer nelas, então tudo bem.

Então, aqui estamos no tipo básico gráfico de dispersão XY com o PIB per capita em um eixo e expectativa de vida no outro. E eles estão coloridos pelo continente, ou região, ou o que seja. Mas se formos para os dados que podemos ver, se você ignorar todo o resto, você pode ver que os valores X estão aqui ligados ao PIB e os valores de Y são a expectativa de vida. E estamos nomeando cada um deles e dando a eles uma região para a cor. Mas isso é realmente algo básico, mesmo se eu me livrasse dessas duas configurações, você pode ver que isso é apenas um gráfico normal de dispersão XY. Vou jogar aqueles de volta lá. Então nós já falamos sobre a cor e coisas assim, então eu não preciso passar por isso de novo, mas este é um gráfico básico de dispersão. Se eu quiser transformar isso em um gráfico de bolhas, preciso adicionar uma terceira dimensão de valor, e essa seria o valor pelo qual as bolhas seriam dimensionadas. Neste caso, esta é a configuração de tamanho de coluna bem aqui, e eu vou colocar essa coluna E. Então vamos dimensionar por esta coluna de população na qual ainda não mexemos até agora. E aí está. Agora podemos ver que as bolhas estão sendo dimensionadas pela população e isso deve aparecer no balão da população agora.

Algumas coisas a saber sobre configuração sobre os pontos: você pode mudar a paleta de cores normalmente aqui, mas você também pode mudar no estilo. Então, eu gosto de fazer baixa opacidade para os pontos, porque, especialmente se você estiver usando uma paleta de cores que é um pouco mais escura, para que possa ver onde estão as sobreposições. Acho que está bem bonito. Você também pode adicionar um contorno aos pontos se quiser. Gosto de manter isso bem mínimo só para dar um pouco de bidimensão a eles. Mas sim, esse é apenas um gráfico de bolhas básico. Há pontos de partida para gráficos básicos de dispersão e bolhas, então você vai conseguir fazê-los tranquilamente.

Outra coisa que eu quero salientar, antes de seguir em frente, sobre esse template é que estamos atualmente no eixo X usando uma escala logarítmica, apenas para apontar que você pode fazer coisas assim neste gráfico.

Ok, o próximo é um gráfico de tira. E isto meio que mostra, é uma boa maneira de exibir dados por categoria e uma outra variável. É algo que as pessoas dos bancos de dados, é um tipo de gráfico que elas realmente gostam de usar. Então temos isso neste template também. A ligação que você vai usar para isso vai ser esta ligação "Series: (connect with line)" nas configurações, mas antes de eu falar sobre isso, devemos falar sobre o que os valores de X e de Y são. Então, para tornar isso possível, você vê que o valor X é esse tipo de valor numérico, e o valor Y vai ser um valor categórico. Então vai ser uma categoria como Europa e Ásia Central. E é isso que faz estas linhas aqui ao longo do eixo Y. E então o valor X coloca os pontos ao longo do eixo. E você vai ver que estes são de fato países e tal. Então é isso que a coluna C está fazendo lá com nome. E depois para fazer estes se conectarem, se eu me livrar disso e ver isso aqui no canto, se eu me livrar disto, ele realmente não faz muita coisa, mas o que ele está fazendo é meio que transformar esses em um categoria e está conectando-os com uma linha. Então isso faz mais sentido se você olhar no gráfico de pontos, na verdade. Mas saiba que gráfico de tira é, basicamente, você tem uma variável categórica e uma variável numérica, e você pode fazer o mesmo mudando a opacidade e tudo com estes. Definitivamente, colocar opacidade menor nestes é realmente muito útil, eu acho, porque há muitas sobreposições interessantes entre os pontos muitas vezes.

Então agora eu vou para o gráfico de pontos. Aqui está outro exemplo. Este é um gráfico muito semelhante. Você tem isso ao longo do eixo X, você tem esse tipo de valor numérico e o eixo Y é a categoria. Mas o eixo Y, cuide de entender, nesse aqui, é também o nome. Então, em vez de ter um tipo de categoria e depois colocar pontos nomeados ao longo dele, como no gráfico de tira, neste temos a categoria e então vamos colocar dois outros pontos de dados ao longo dela para aquela mesma categoria. Então neste caso temos o país e o ano e a renda naquele ano, e estamos conectando-os com o ponto, se isso faz sentido. Então 2014 e 2004. E certifique-se de ter esta série. Neste caso, não estamos realmente conectando pelo valor como fizemos no outro. Estamos conectando no país. Se você quiser que isso seja por país, cuide para que essa seja a coluna que você tem nesta configuração "Series (connect with line)". Então, se eu me livrar disso, ele remove as linhas. Ele só tem os pontos colocados ao longo, mais ou menos como nós tivemos no gráfico de tira. Se colocamos de volta, ele meio que traça a linha entre eles. E há várias opções de estilo. Há vários tipos de opções de estilo para a linha e coisas assim.

A penúltima que eu vou mostrar é a opção de distribuição, e estes são os diagramas de caixa, violino, e "beeswarm". Basicamente, o que você precisa fazer para estes é introduzir valores de X e de Y. Seu valor de X é mais provável que seja uma categoria mais uma vez, como você pode ver. E os valores Y serão como o valor de, a partir de cada uma dessas categorias, algum tipo de valor. E estes são ótimos para mostrar a distribuição de valores em diferentes categorias. Assim, diferentemente dos diagramas de tiras e pontos, você geralmente vai ter suas variáveis categóricas ao longo do X e seus valores, seus valores numéricos, vão ser o Y. E se você for no painel de configurações deste diagrama de caixa, violino ou beeswarms, estes são todos diferentes, mais um tipo de opções estatísticas, opções de visualização que as pessoas da ciência e das estatísticas gostam de usar para mostrar distribuição. Portanto, temos beeswarm, que é quando você agrupa os valores como um enxame de abelhas. E então você também tem, você pode ver que eles estão na verdade em cima uns dos outros aqui, e quando você agrupa, eles se espalham e ficam com a aparência de um enxame de abelhas.

E então você também pode ter, também podemos ter o diagrama de caixa para mostrar a distribuição da mediana, e você também pode ter o diagrama de violino, que é semelhante ao beeswarm, se você seleciona ele, ele meio que contorna o beeswarm, basicamente.

Estas são opções avançadas e não são algo que muitas pessoas usam, mas se você curte ciência e estatística, eles são definitivamente úteis e eles estão lá para você usar.

Finalmente, eu só quero mostrar isso porque eu não vou realmente explicar, é meio complicado. Mas também temos esse Hans Rosling, que foi um famoso estatístico, economista, e este gráfico faz uma animação ao longo de diferentes anos e mostra diferentes pontos de desenvolvimento de diferentes países do mundo. E esta é uma opção. Você faz muita coisa aqui neste controle deslizante de tempo. E há essa configuração extra de coluna de tempo aqui que você tem que usar. Não vou realmente mostrar isso agora, porque se você curtir isso, você vai ser capaz de entender sozinho.

Mas este é um bom lugar para encerrar. Este provavelmente será o vídeo mais longo, porque este é o nosso template mais versátil, mas uma vez que você sabe o que ele pode fazer, é realmente muito poderoso. Então obrigada por ouvir.