

## La plantilla de dispersión

En este video, hablaré un poco más en profundidad sobre la plantilla de dispersión, y esta es una plantilla que posiblemente sea la más amplia en sus usos que cualquier otra plantilla de Flourish. Siempre estamos bromeando cuando estamos desarrollando una nueva plantilla decimos como “oh, tal vez podamos agregar eso al gráfico de dispersión”, porque puede hacer tantas cosas.

Y quería comenzar aquí en el selector de plantillas, porque quería mostrarles visualmente cuántas cosas diferentes puede hacer la plantilla de dispersión. Por lo tanto, estos son todos los puntos de partida que tenemos con la dispersión, y estos son todos los diferentes tipos de gráficos, generalmente de hecho, diferentes tipos de gráficos que pueden hacer con la plantilla de dispersión. Creo que el único que no está aquí y que pueden hacer es el diagrama “beeswarm”. Así que voy a ir por ellos y explicaré la estructura básica de la plantilla de dispersión y luego mostraré cómo se hacen cada uno de los otros diagramas. Pero sepan que si desean hacer uno de estos, no tienen que hacerlo al 100%, con solo ingresar sus datos en este gráfico básico, luego en el gráfico de dispersión y luego intentar realizar los ajustes ustedes mismo en la configuración de la columna. Existen estos puntos de partida que deberían ser buenas guías.

Así que comenzaré aquí con un diagrama de dispersión básico que muestra, creo, el PIB per cápita y la expectativa de vida de varios países. Cada uno de estos puntos es un país. Y quería mostrar esto porque quiero mostrarles las diferencias entre un gráfico de dispersión básico y un gráfico de burbujas. Aquí está el diagrama de dispersión básico, también es bueno saber que en la plantilla de dispersión, el panel de configuración es algo enorme y un poco intimidante. Pero no se asusten por eso. Muchas de estas cosas, si no están haciendo este tipo de gráfico, no tienen que tocarlas, así que está bien.

Así que aquí estamos en el tipo de diagrama de dispersión XY básico con el PIB per cápita en un eje y la expectativa de vida en el otro. Y están coloreados por continente, o región, o lo que sea. Pero si vamos a los datos que podemos ver, si ustedes simplemente ignoran todo lo demás, pueden ver que los valores X están aquí vinculados al PIB y los valores Y representan la expectativa de vida. Y estamos nombrando a cada uno de ellos y dándoles una región para el color. Pero esto es realmente básico, incluso si me deshago de estas dos configuraciones, pueden ver que esto es solo un diagrama de dispersión XY normal. Los devolveré allí, sin embargo. Ya hemos hablado sobre el color y cosas por el estilo, así que no tengo que repasar eso, pero este es un diagrama de dispersión básico. Si quiero convertir esto en un gráfico de burbujas, necesito agregar una tercera dimensión de valor, y ese sería el valor por el cual se dimensionarían las burbujas. En este caso, esta es la configuración de tamaño (size) de la columna aquí, y voy a hacer esta columna E. Así que vamos a dimensionarlo por esta columna de población que aún no hemos tocado. Y ahí está. Ahora podemos ver que las burbujas están siendo dimensionadas por la población y esto debería aparecer en el globo de población ahora.

Algunas cosas que deben saber sobre la configuración de los puntos: aquí pueden cambiar la paleta de colores normalmente, pero también pueden cambiar el estilo. A mí me gusta hacer una opacidad baja para los puntos, porque especialmente si están usando una paleta de colores que es un poco más oscura, puedan ver dónde se superponen. Creo que luce muy bien. También pueden agregar un contorno a los puntos si lo desean. Me gusta mantenerlo al mínimo solo para darles un poco de dimensión. Pero sí, esto es sólo un gráfico de burbujas básico. Estos son puntos de partida para la dispersión básica y los gráficos de burbujas, por lo que lo podrán hacer sin problemas.

Otra cosa que quiero señalar, antes de continuar, sobre esta plantilla es que actualmente estamos en el eje X usando una escala logarítmica, “log scale” solamente para mostrarles que pueden hacer cosas como esta en este gráfico.

Bien, el siguiente es un gráfico de franjas. Y este tipo de presentación es una buena forma de mostrar datos por categoría y otra variable. Es algo que las personas en las bases de datos, es un tipo de gráfico que realmente les gusta usar. Así que también tenemos esto en esta plantilla. El enlace que usarán para esto será este enlace “Series: (connect with line)” en la configuración, pero antes de hablar sobre eso, deberíamos hablar sobre cuáles son los valores del eje X y eje Y. Para que esto sea posible, verán que el valor X es ese tipo de valor numérico, y el valor Y será un valor categórico. Entonces será una categoría como Europa y Asia Central. Y eso es lo que hace estas líneas aquí a lo largo del eje Y. Y luego el valor X coloca los puntos a lo largo de esto. Y verán que estos son países y cosas así. Entonces, esto es lo que la columna C está haciendo allí con el nombre. Y luego, para hacer que se conecten, si me deshago de esto y lo veo aquí en la esquina, si me deshago de esto, realmente no hace mucho, sino lo que está haciendo es convertirlos en una categoría y se conecta ellos con una línea. Entonces eso tiene más sentido si nos fijamos en el gráfico de puntos, de hecho. Pero tengan en cuenta que el gráfico de franjas es básicamente que tienen una variable categórica y una variable numérica, y pueden hacer lo mismo cambiando la opacidad y todo con estas. Definitivamente, poner una opacidad más baja en estos es realmente muy útil, creo, porque a menudo hay muchas superposiciones interesantes entre los puntos.

Así que ahora voy al gráfico de puntos. Aquí hay otro ejemplo. Este es un gráfico muy similar. Entonces tienen esto a lo largo del eje X, tienen este tipo de valor numérico y el eje Y es la categoría. Pero el eje Y, asegúrense de entender, en este caso, también es el nombre. Entonces, en lugar de tener un tipo de categoría y luego colocar puntos con nombre en ella, como en el gráfico de franjas, en este tenemos la categoría y entonces colocamos otros dos puntos de datos para esa misma categoría. Entonces, en este caso tenemos el país y luego tenemos el año y los ingresos de ese año, y los estamos conectando al punto, si eso tiene sentido. Así que 2014 y 2004. Y asegúrense de tener esta serie. En este caso, no estamos realmente conectando por valor como lo hicimos en el otro. Nos estamos conectando en el país. Si desean que esto sea por país, asegúrense de que esta sea la columna que tiene en esta configuración de “Series (connect with line)”. Entonces, si me deshago de él, elimina las líneas. Sólo tiene los puntos colocados a lo largo, más o menos como los que teníamos en el

gráfico de franjas. Si lo volvemos a poner, dibuja la línea entre ellos. Y hay todo tipo de opciones de estilo. Hay varios tipos de opciones de estilo para la línea y cosas así.

El penúltimo que les mostraré es la opción de distribución, y estos son los diagramas de caja, violín y “beeswarm”. Básicamente, lo que necesitan hacer para esto es ingresar los valores de X y de Y. Sus valores de X es más probable que sean una categoría otra vez, como pueden ver. Y los valores Y serán como el valor, de cada una de estas categorías, algún tipo de valor. Y estos [gráficos] son excelentes para mostrar la distribución de valores en diferentes categorías. Entonces, a diferencia de los diagramas de franjas y puntos, generalmente tendrán sus variables categóricas a lo largo del X y sus valores, sus valores numéricos, serán el Y. Y si van al panel de configuración de este diagrama de caja, violín y beeswarm, estos son todos diferentes, más como un tipo de opciones estadísticas, opciones de visualización que a las personas en ciencias y estadísticas les gusta usar para mostrar la distribución. Entonces tenemos beeswarm, que es cuando ustedes agrupan sus valores como un enjambre de abejas. Y luego ustedes también tienen, pueden ver que de hecho ellos están en verdad uno encima de otro aquí, y cuando activan la opción beeswarm, estos se extienden y se ven como un enjambre de abejas.

Y luego pueden tener también, también podemos tener el diagrama de caja para mostrar la distribución de la mediana, y también pueden tener el diagrama de violín, que es similar al beeswarm, si lo seleccionan, es como un esquema de beeswarm, básicamente.

Estas son opciones avanzadas y no son algo que muchas personas usan, pero si disfrutan de la ciencia y la estadística, definitivamente son útiles y están ahí para que las usen.

Finalmente, solo quiero mostrarles esto porque realmente no lo explicaré, es un poco complicado. Pero también tenemos a este Hans Rosling, que era un famoso estadista, economista, y este gráfico anima durante diferentes años y muestra diferentes puntos del desarrollo de diferentes países del mundo. Y esta es una opción. Hacen mucho aquí en este control deslizante de tiempo. Y hay una configuración de columna de tiempo adicional aquí que tienen que usar. Realmente no lo mostraré ahora, porque si les interesa esto, podrán entenderlo por ustedes mismos.

Pero este es un buen lugar para terminar. Este será probablemente el video más largo, porque esta es nuestra plantilla más versátil, pero una vez que saben lo que pueden hacer, es realmente muy poderosa. Así que gracias por escuchar.