Lab 8. Tableau Visualization

In this lab, we will cover the following topics:

- Dual axis waterfall charts
- Pareto charts
- Bump charts
- Sparkline charts
- Donut charts
- Motion charts

Technical requirements

We will use Tableau 2019.x and datasets about potholes, coal emissions, soccer rankings, alien sightings, avocado prices, and snowfall.

Introduction

We will go beyond **`Show Me`** and learn techniques to master visualizations in Tableau to make your dashboards stand out. In this recipe, we will cover detailed steps for creating visualizations that convey percentages, based on how a proportion relates to a whole, or how a value changes. The use cases vary from identifying elements in the data that create the biggest impact, creating ranks for different categories over a period of time, or tracking goals for organizations.

Dual axis waterfall chart

Waterfall charts show us how a series of positive or negative changes contribute to a final total. Waterfall charts can also be used to show the composition of a measure with categorical values, resulting in the cumulative amount. The **`Dual Axis`** waterfall charts illustrate changes more elegantly than color alone.

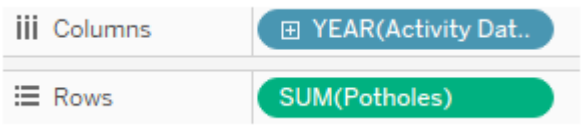
Getting ready

In this recipe, we will create a dual axis waterfall chart, using a line chart as the base.

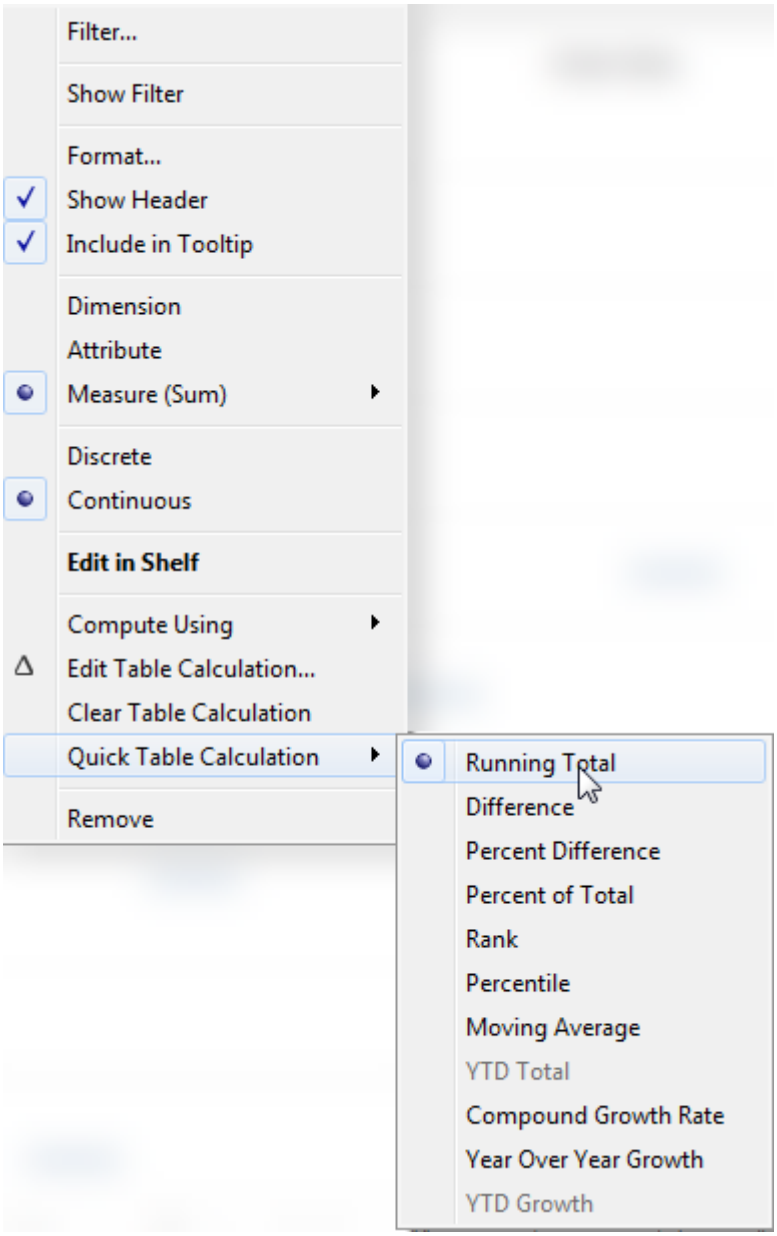
How to do it..

We're using data from the city of Chicago to see how many potholes are identified and how many are fixed. Use `Waterfall.tbwx` and `Potholes.csv` to follow along:

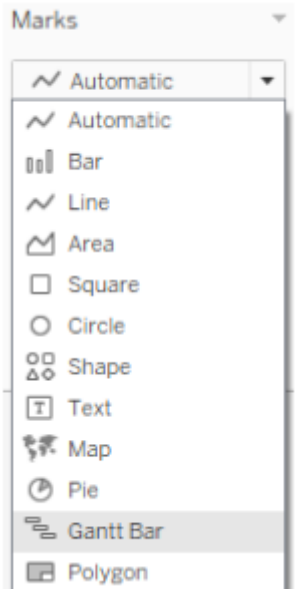
1. Create a line or bar chart with one dimension and one measure---**`Activity Date`** and **`Potholes`** as follows:



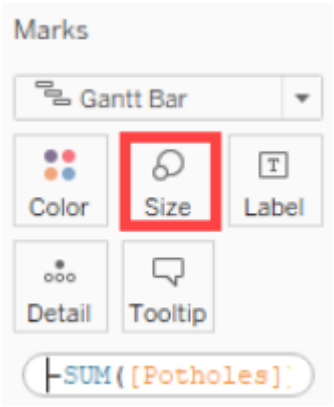
2. Apply a **`Running Total`** measure under **`Quick Table Calculation`** for **`Potholes`**, as shown in the following screenshot:



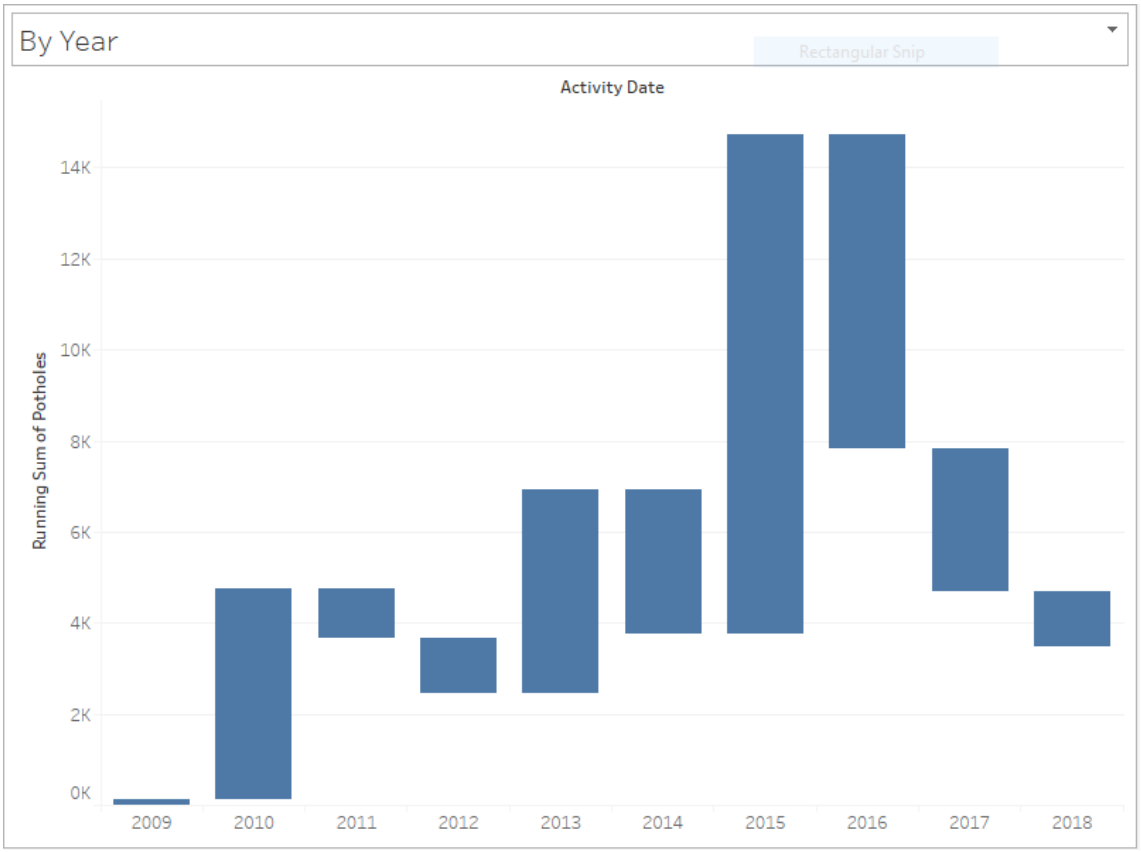
3. Change the `Marks` type to `Gantt Bar`, as shown in the following screenshot:



4. Create bars by sizing the mark with a negative version of the same measure. Sizing by a negative version allows the start of each bar to line up with the end of the previous bar. We illustrate one method by applying negative **`Potholes`** to **`Size,`** as shown in the following screenshot:



5. At this point, we should have a single axis waterfall chart, as follows:

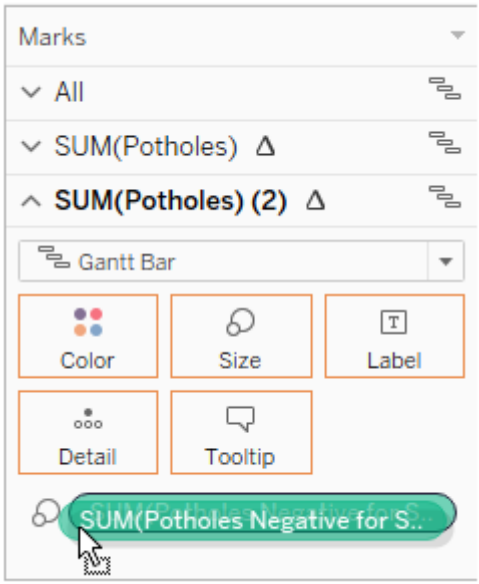


Note

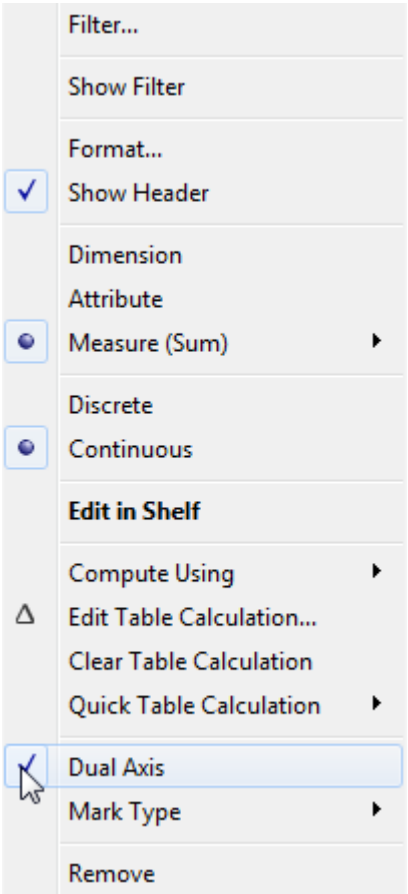
Waterfall charts can be used with time and categorical dimensions. The order of the categories can impact the story. Color can be used to show negative and positive values. We can have more options with shapes using a dual axis waterfall chart.

6. Duplicate the pill by pressing the [*Ctrl*] key and selecting **`SUM(Potholes)`** and placing to the right, as shown in the following screenshot:

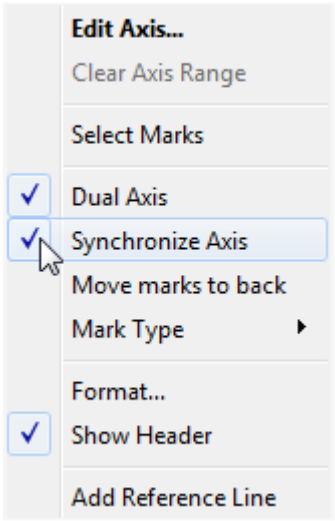
7. Clear the size from the **`Marks`** card for the second measure, as shown in the following screenshot:



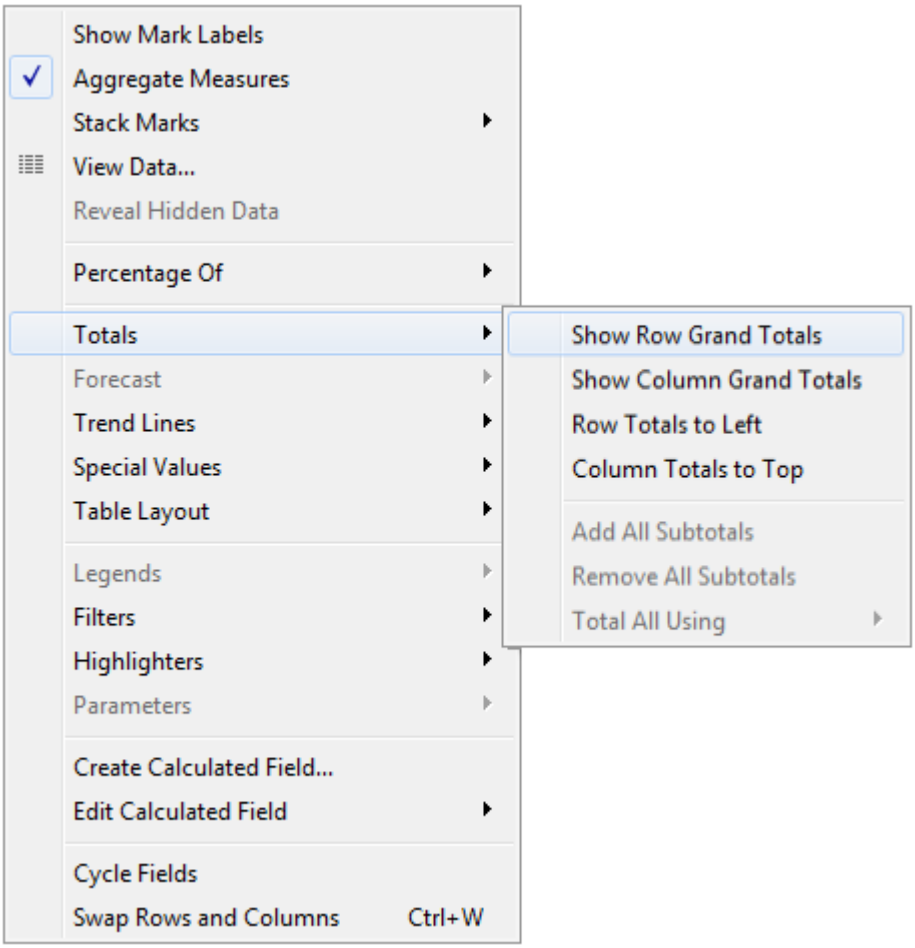
8. Right-click on the second green pill and choose `Dual Axis`, as shown in the following screenshot:



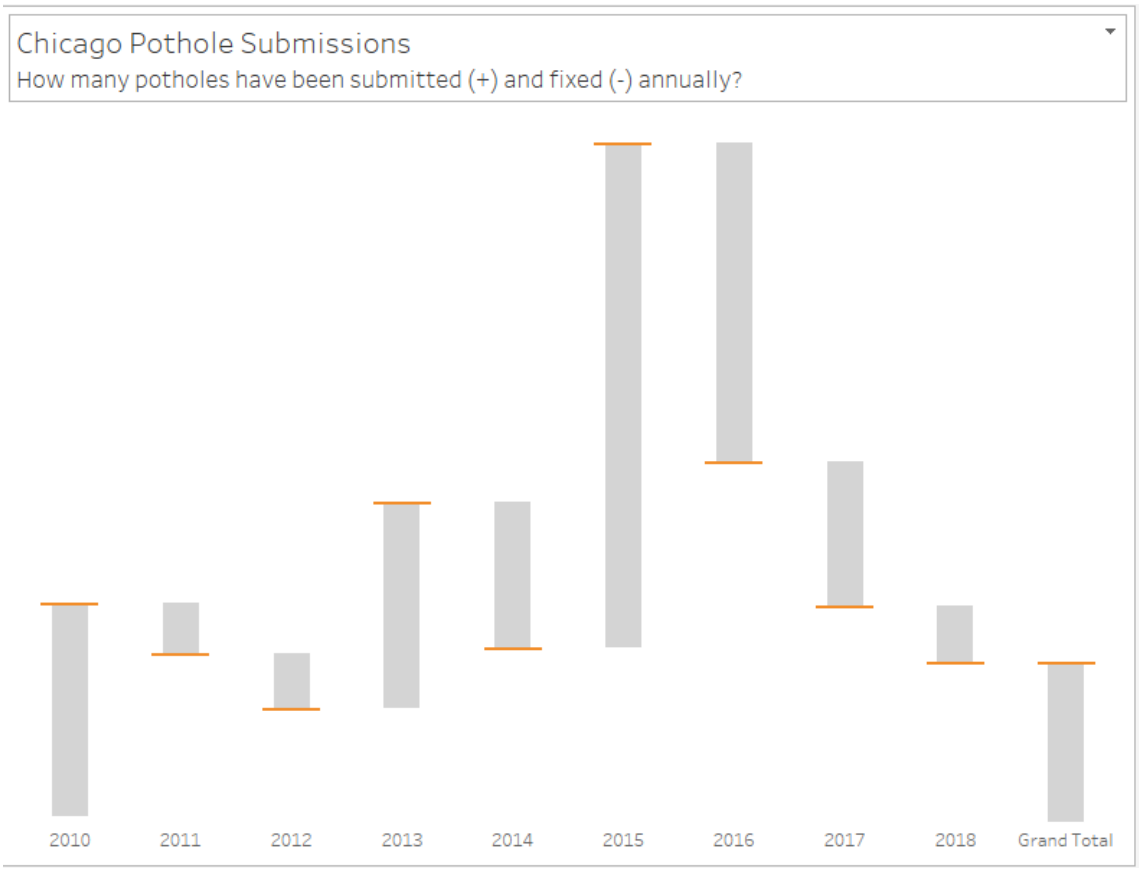
9. Click on the second axis, select `Synchronize Axis`, and hide the headers, as shown in the following screenshot:



10. Add **`Totals`** from the **`Analysis`** menu. Navigate to ** `Totals` **| **`Show Row Grand Totals`**, as shown in the following screenshot:



11. In the `Marks` card, change the default mark or modify the color of the mark based on whether a value is negative or positive. We use `Gantt Bar` to create lines at the end of each bar, as shown in the following screenshot:



How it works...

In [*step 1*], we create a line chart with one dimension and one measure and then apply `Running Total Table Calculation` in [*step 2*]. In [*step 3*], we then change the mark type to `Gantt Bar`. In order to create bars, in [*step 4*], we negate the measure so each column can grow to the previous mark. This creates the stair step effect, so the end of one mark is the beginning of the previous one. We see in [*step 5*] that we've created a basic waterfall chart.

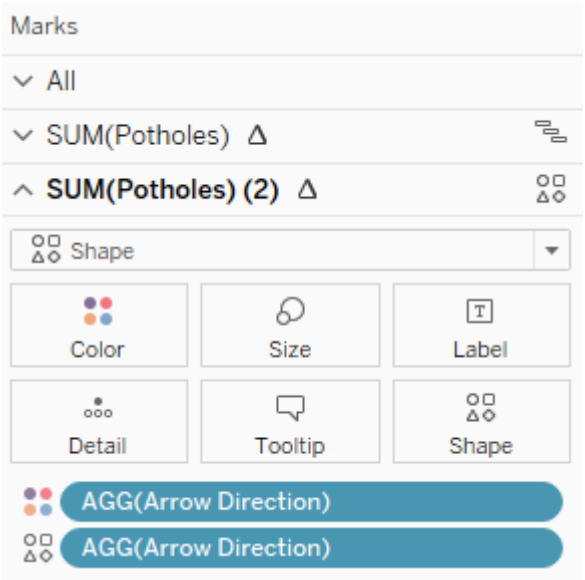
In order to highlight the positive or negative change, we create a `Dual Axis` chart. In [*step 6*], we duplicate the same measure on the `Rows` shelf. We clear the size from the second `Marks` card in [*step 7*], which will emphasize the ending value when we make the dual axis in [*step 8*]---causing the marks to overlay. In [*step 10*], we add a total to show the cumulative impact of all values. Finally, in [*step 11*], we change the default mark type to `Gantt` Bar to emphasize the negative or positive change.

There's more...

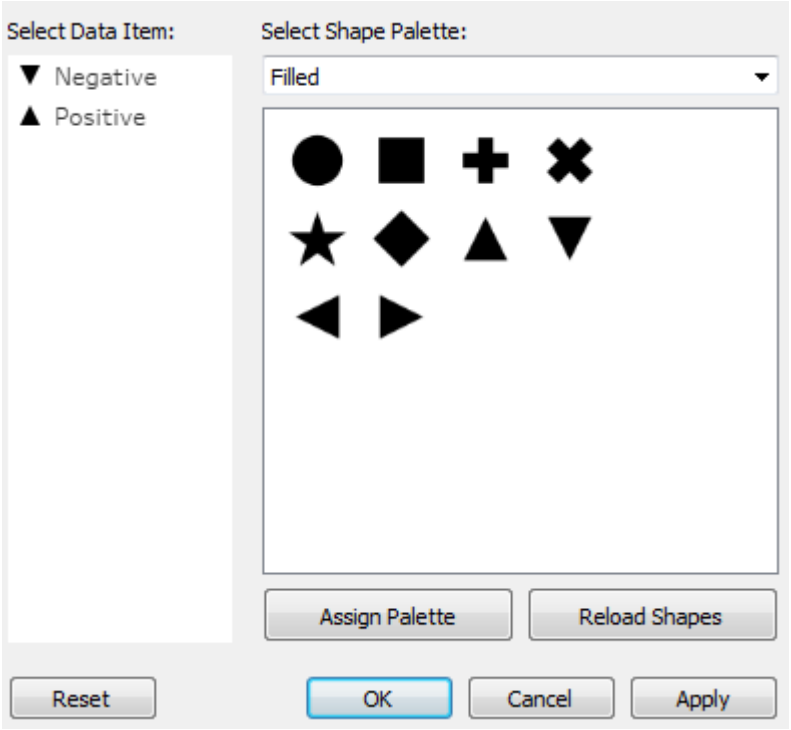
We can also use shapes to indicate positive or negative values by taking the following steps.

1. Create calculations for positive and negative values, as follows:

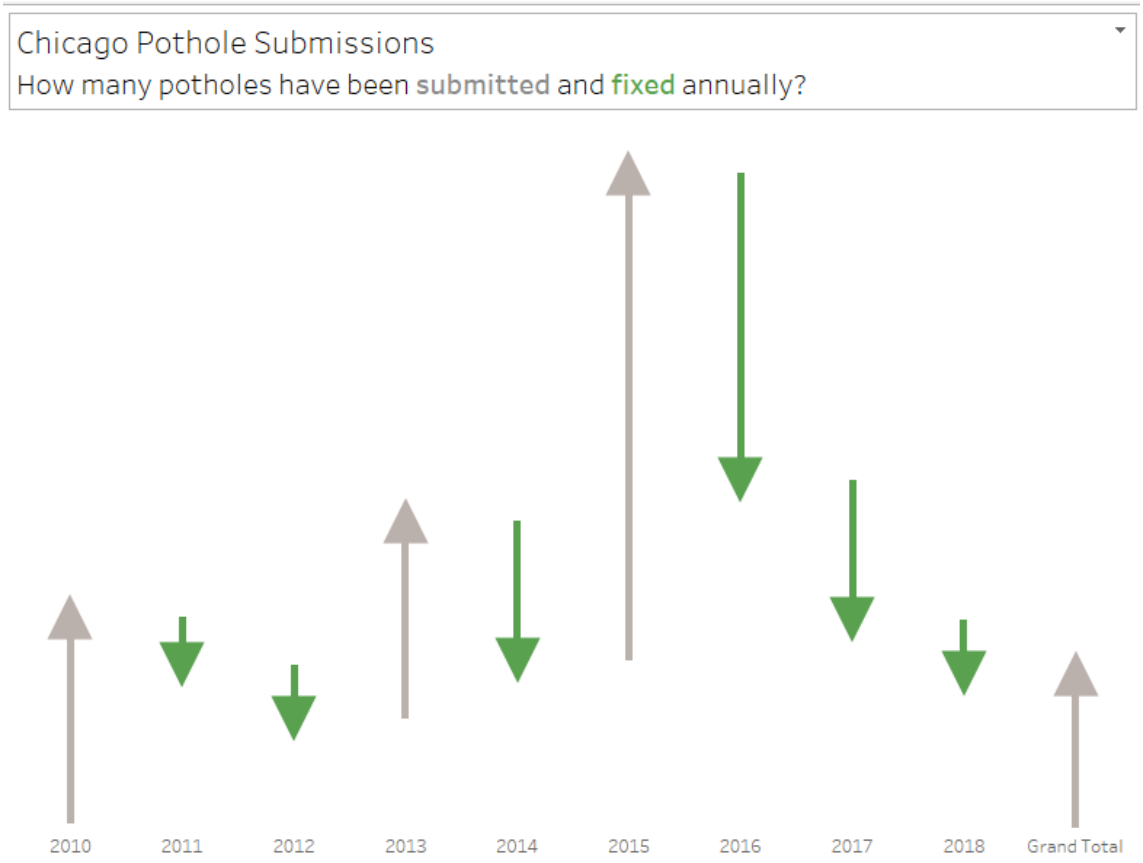
2. Apply **`Arrow Direction`** to **`Shape`** (and optionally **`Color`** for both **`SUM(Potholes)`** and **`SUM(Potholes) (2)`**, as follows:



3. Click on **`Shape`**, select **`Filled`**, and choose the up and down triangles, as follows:



Our final visualization can be seen in the following screenshot:



Pareto chart

Pareto charts are best used when you want to identify the biggest contributing factors to a measurable outcome. They help with identifying priorities and are often used in risk management. The name comes from Vilfredo Pareto, the originator of the 80/20 Rule, or the Pareto Principle, which roughly states that 80% value comes from 20% of a factor. An example would be 80% of sales comes from 20% of our customers.

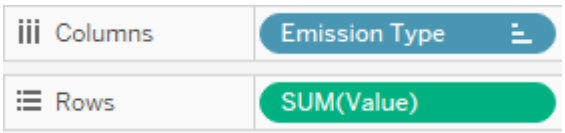
Getting ready

In this section, we will use a bar chart with reference lines to create a Pareto chart.

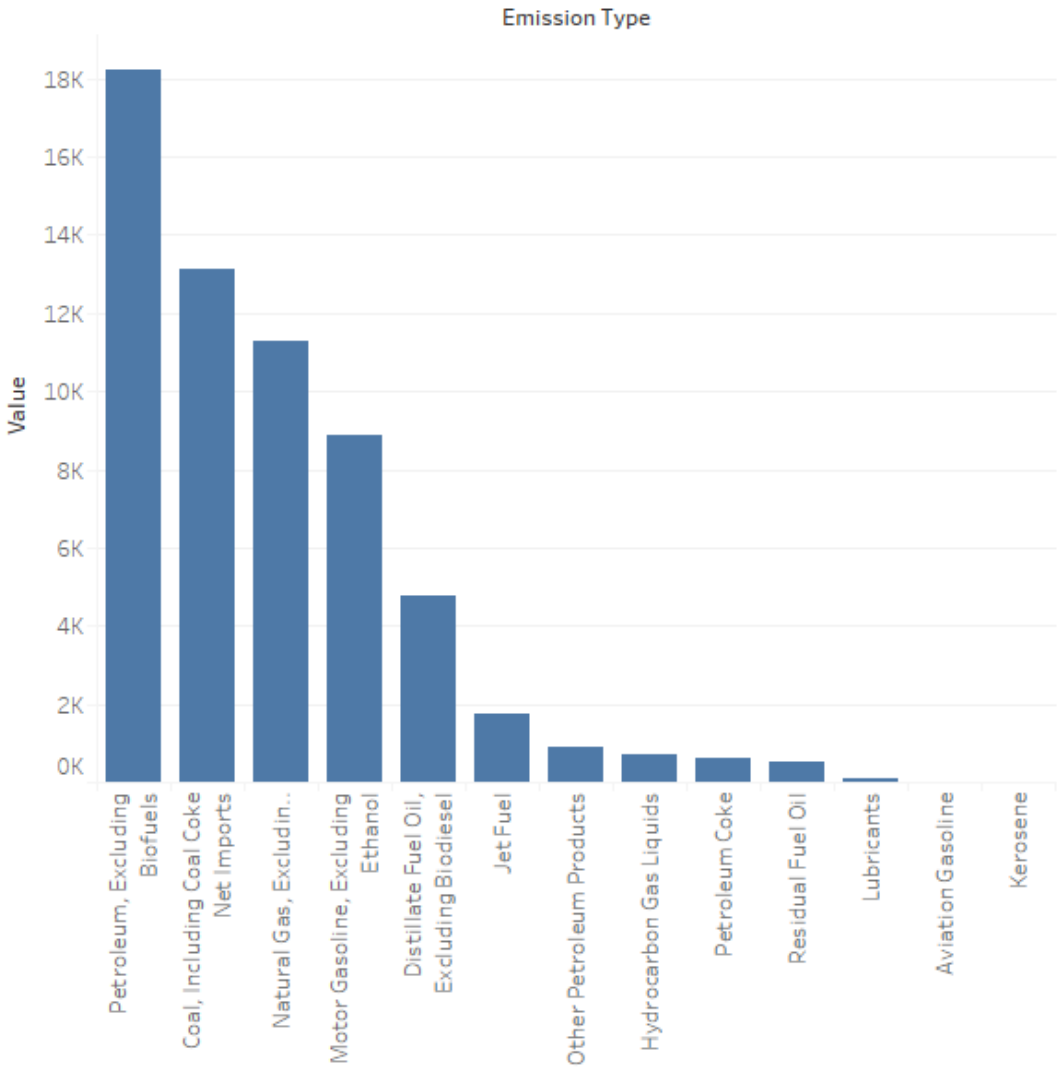
How to do it..

In this example, we are looking at CO2 emissions data. Open the packaged workbook `Pareto.twbx` to follow along:

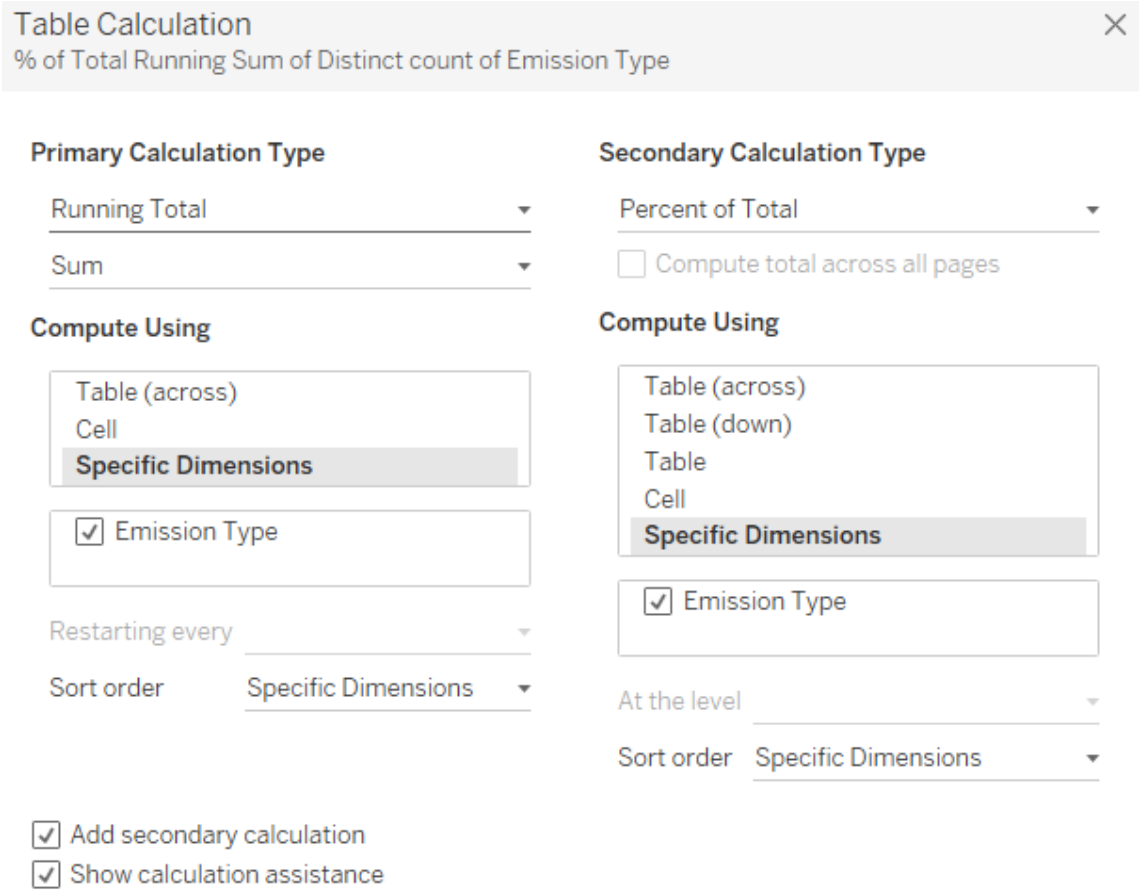
1. Create a bar chart with **`Emission Type`** on **`Columns`** and **`Value`** on **`Rows`**, as follows:



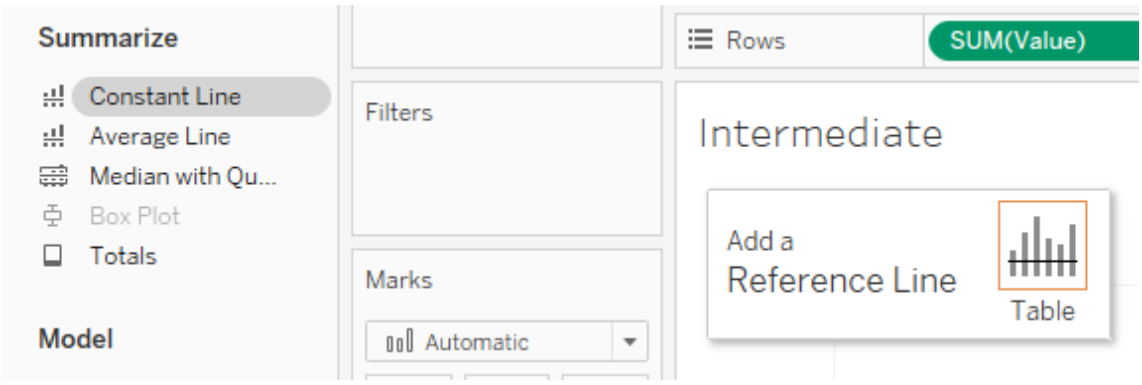
2. Order the chart in descending order according to the measure, as follows:



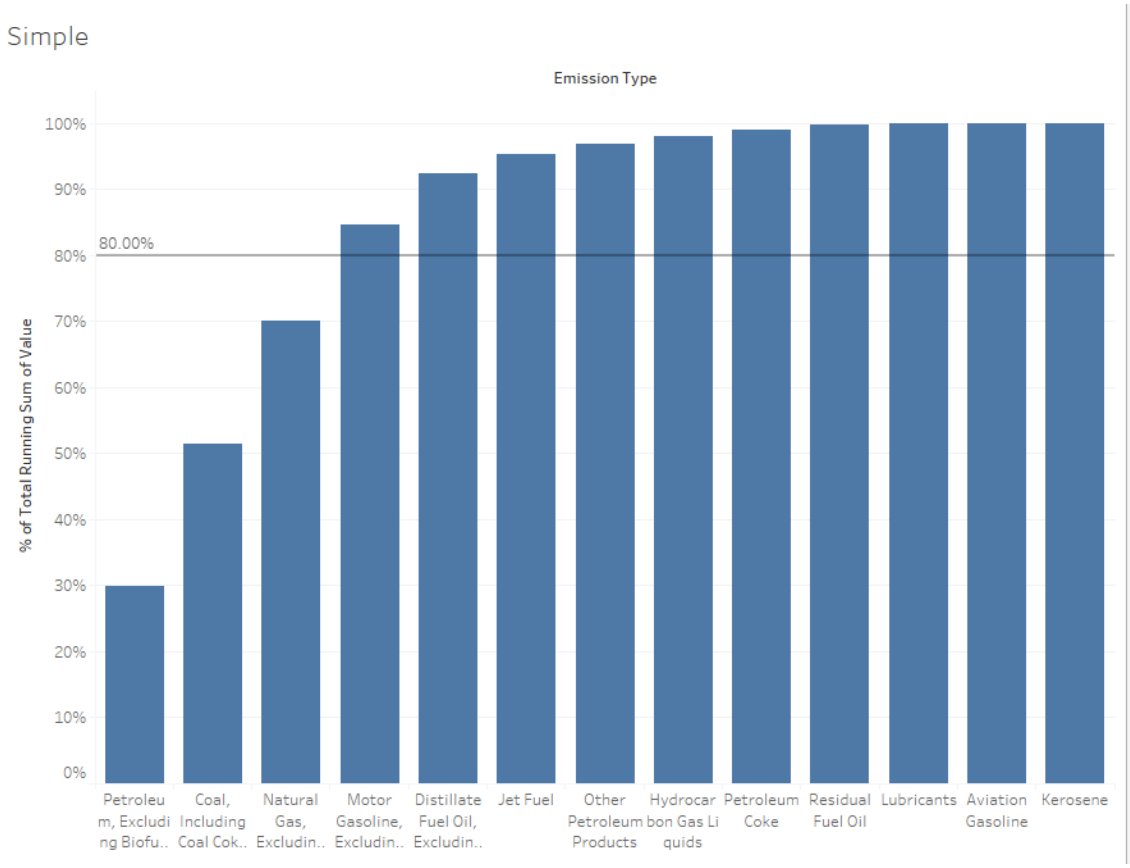
3. Add a primary and secondary table calculation for **`Running Total`** and **`Percent of Total`**. Modify **`Compute Using`** to **`Specific Dimensions`**, as depicted in the following screenshot:



4. Add a **`Reference Line`** at 80%, as follows:



In the following screenshot, we can see our basic **`Pareto`** chart:



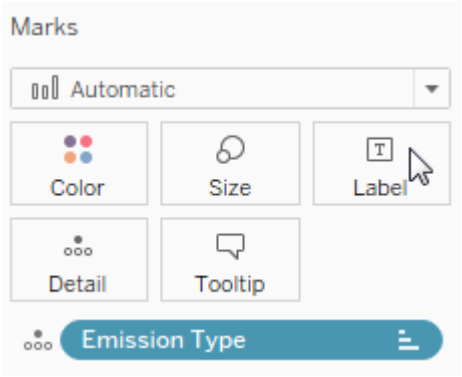
How it works...

In [*step 1*] and [*step 2*], we create a bar chart using one dimension and one measure sorted in descending order. For [*step 3*], we apply the `Running Total` and `Percent of Total Table Calculation`. In [*step 4*], we use reference lines at 80% and a simple bar chart to allow us to see what categories have the biggest impact.

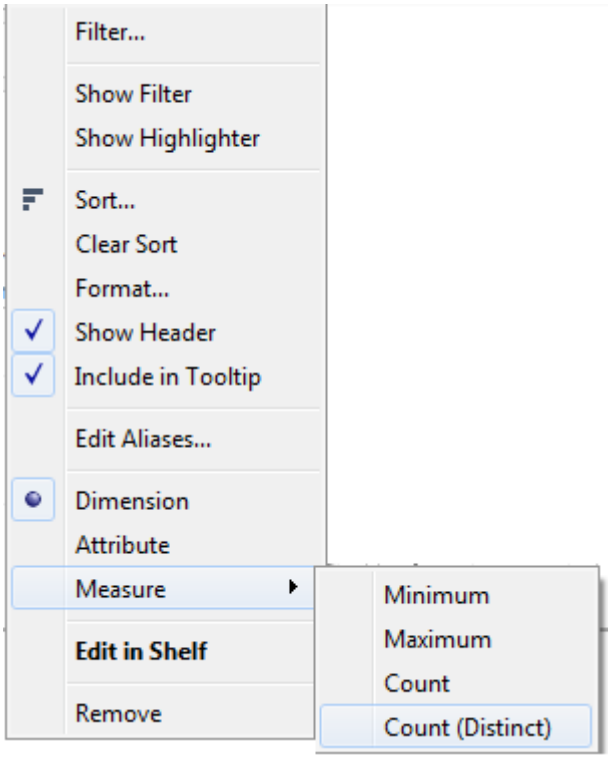
There's more...

We can also create a more advanced version of the Pareto chart that shows the actual 80% and 20% values. We can start with the simple version created in the previous steps, and use the advanced worksheet to see the end product, by taking the following steps.

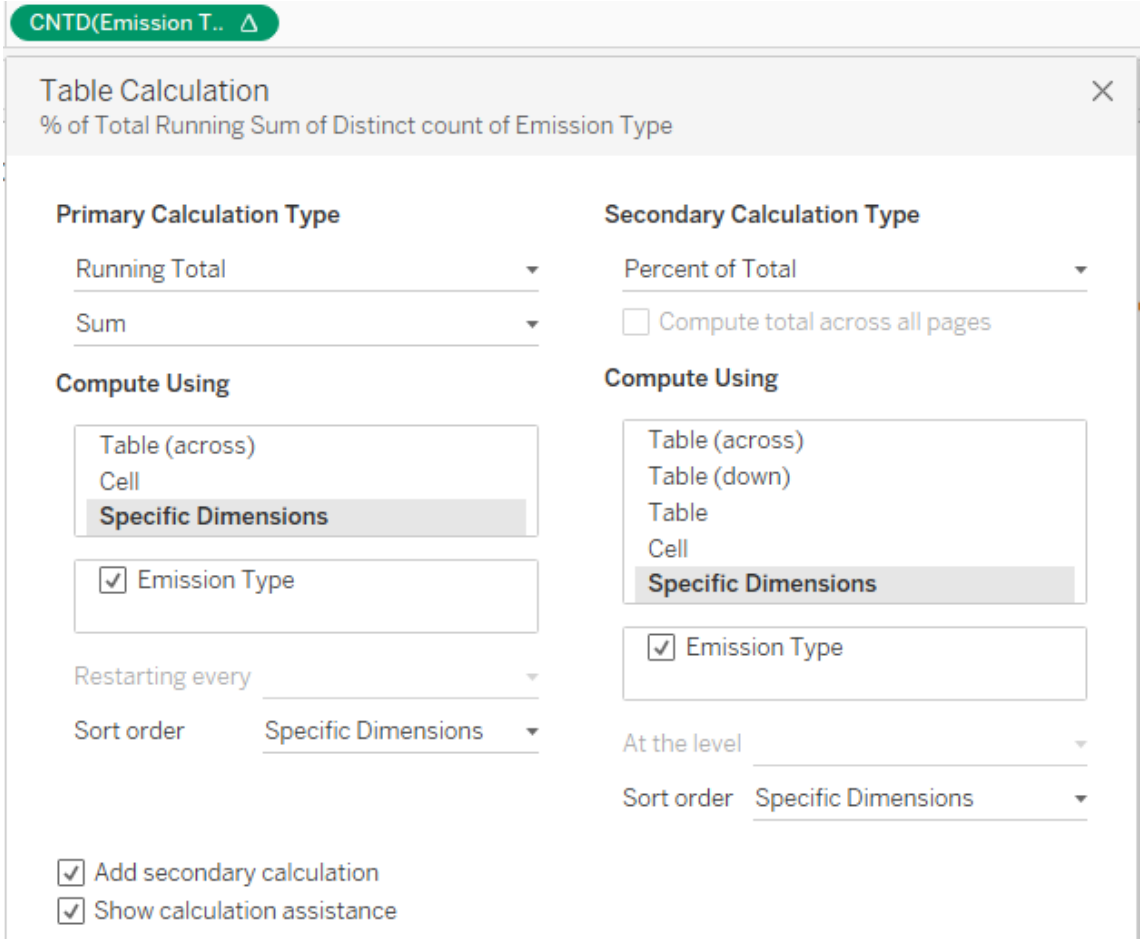
1. Add `Emission Type` to the `Detail` section of the `Marks` card. We do this because our table calculation for `Sum(Value)` is referencing a specific dimension `Emission Type`, as follows:



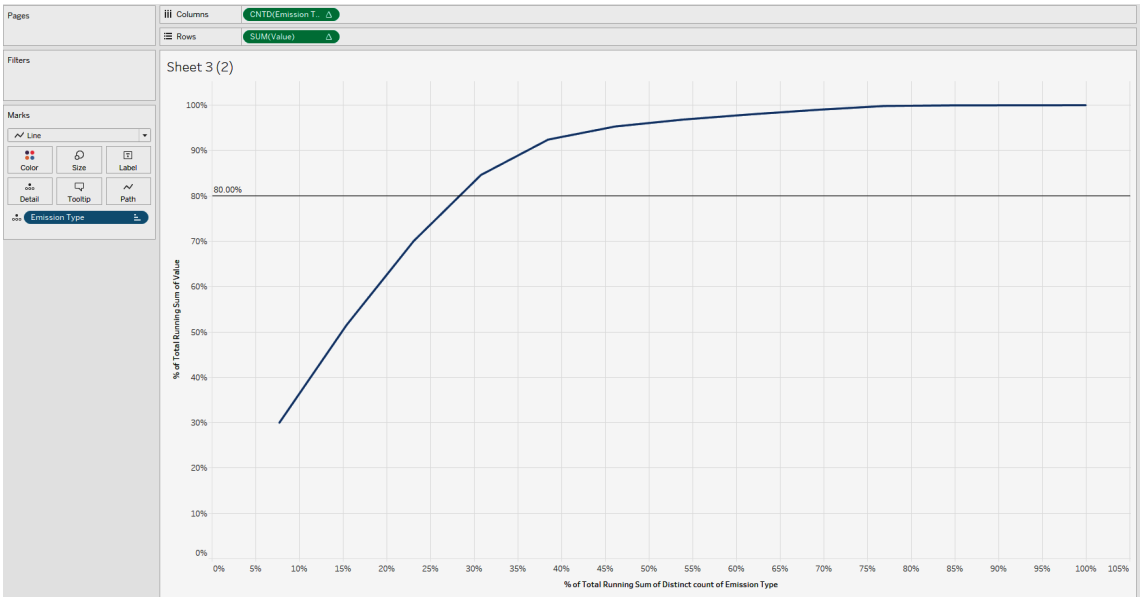
2. On the `Columns` shelf, modify `Emission Type` to be a `Count (Distinct)`, as follows:



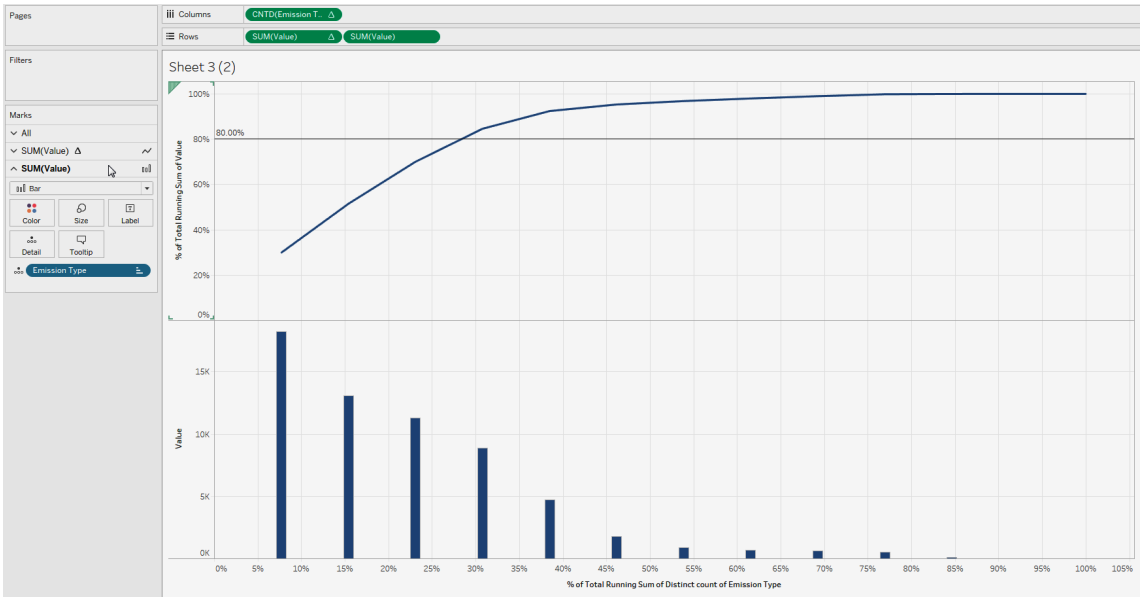
3. Apply the `Running Total` and `Percent of Total` table calculation to it, as follows:



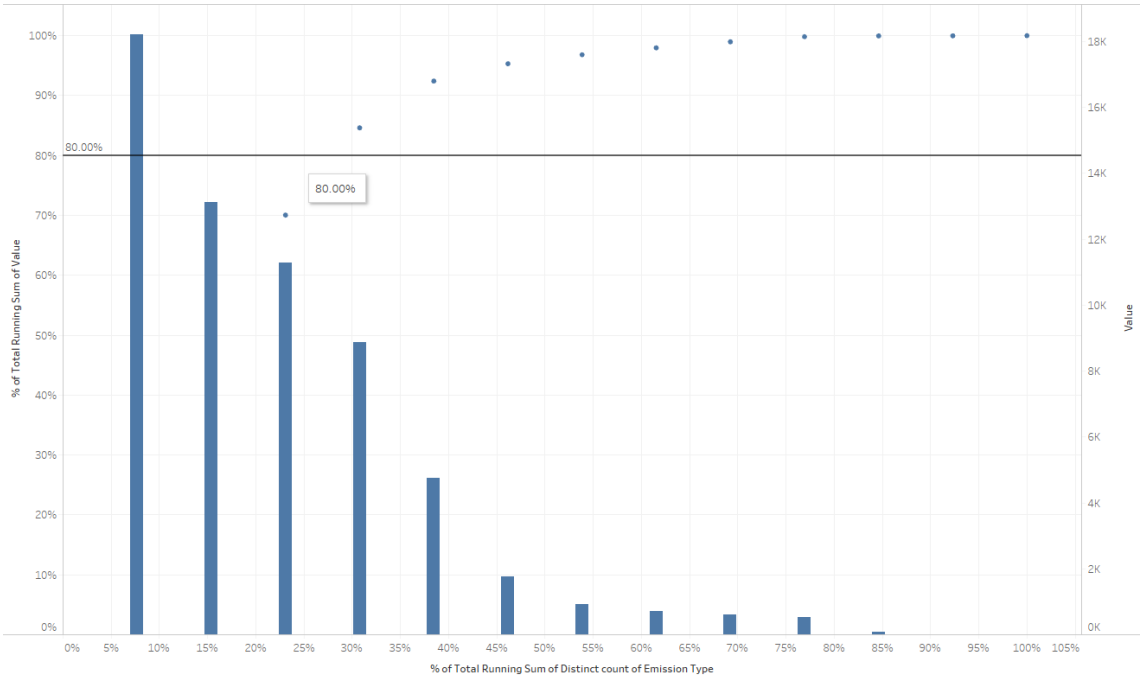
4. Change the **`Mark`** type to **`Line`**, as follows:



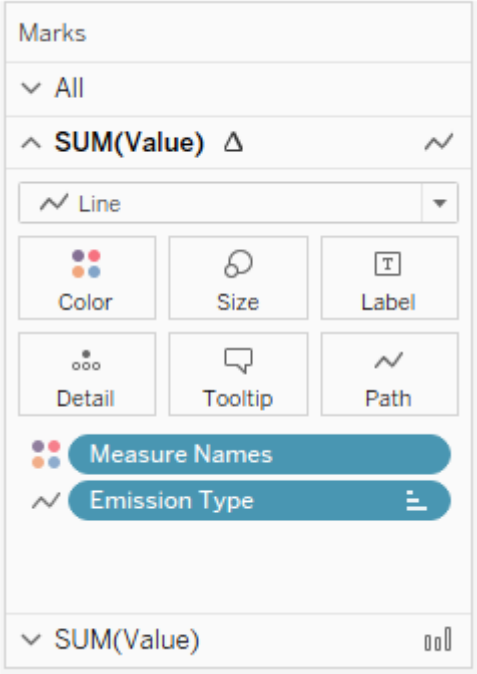
5. Add value to **`Rows`** to begin making a **`Dual Axis`** chart and change the mark type to **`Bar`**, as follows:



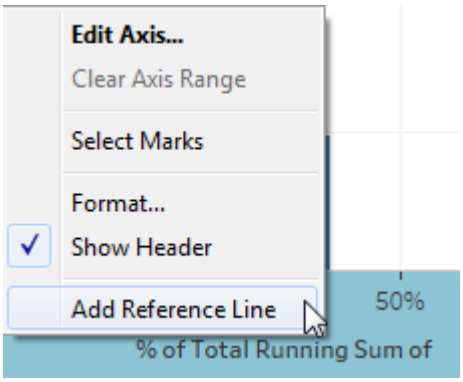
6. When you switch the newly placed green pill to dual axis, it'll change the line chart to a scatter plot as a result of having essentially a measure in the columns and rows, as shown in the following screenshot:



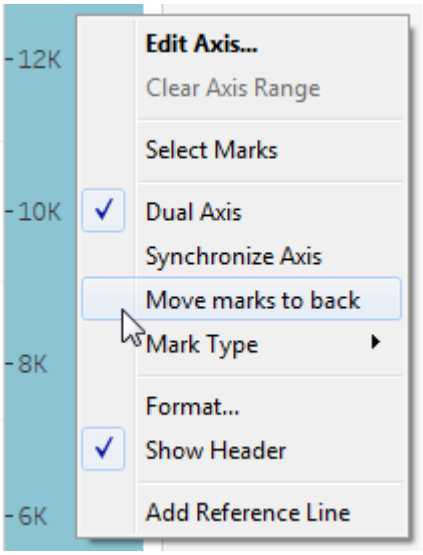
7. Change the dots to a continuous line, and move `Emission Type` to `Path`. Make sure you are doing this for the first green pill, as follows:



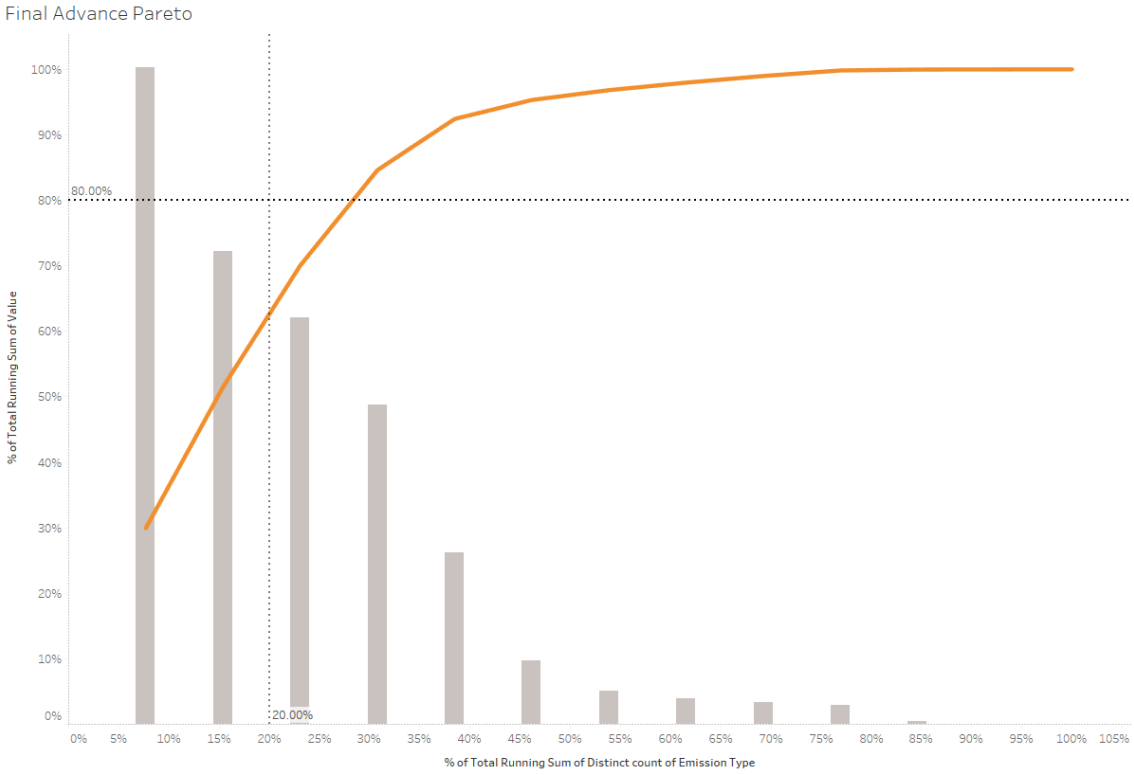
8. Add a reference line for 20%, as follows:



9. Right-click on the axis and select the option **`Move marks to back`**, which allows the reference lines to sit on top of the visualizations, as follows:



After some cleanup, it should look like the following screenshot:



Bump chart

We will use the bump chart to compare two dimensions against each other using a single value of measurement. They are great for examining rank changes.

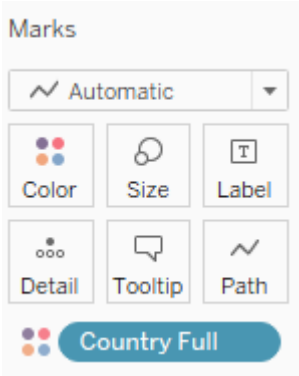
Getting ready

For the bump chart recipe, we will use a line chart with the rank table calculation.

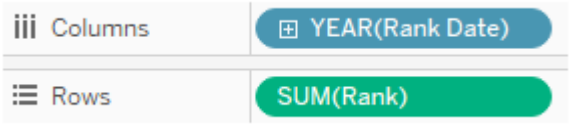
How to do it..

In this example, we are ranking the teams of the South American Soccer Confederation. Use the `Bump Chart.twbx` workbook to follow along:

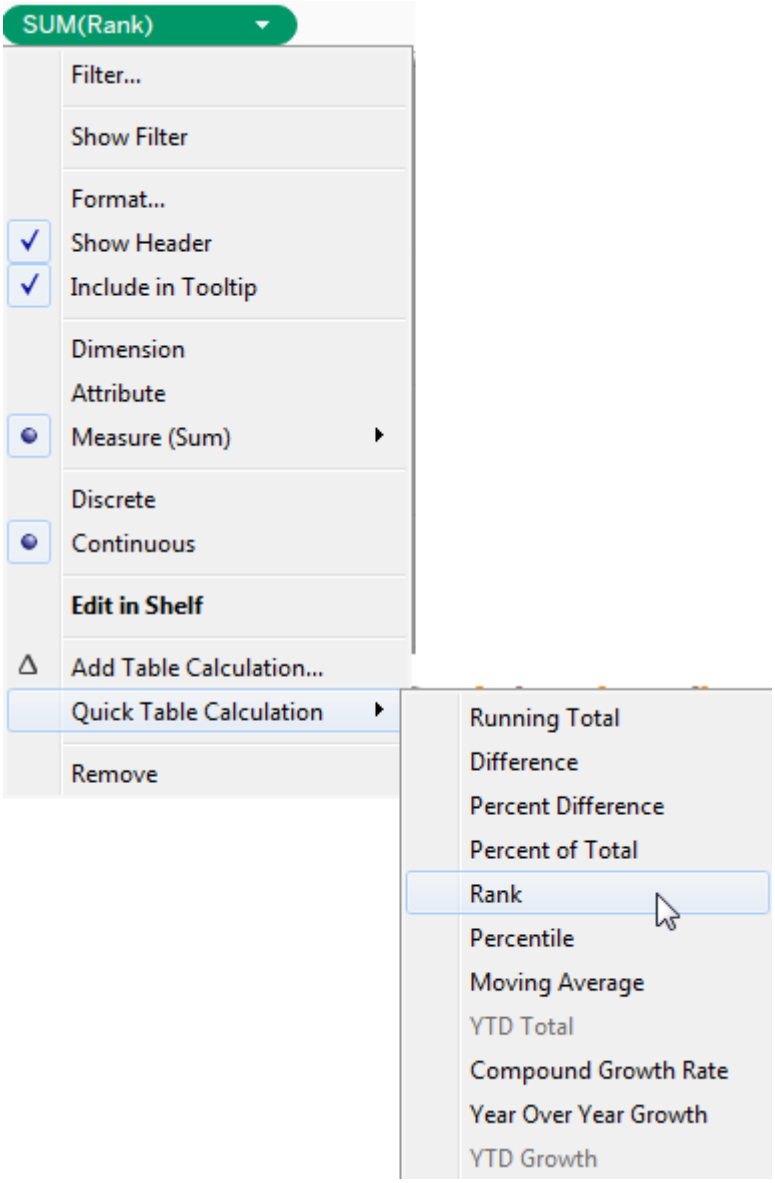
1. Since we want to color our visualization lines by country, put **`Country Full`** on [**`Color`**], as follows:



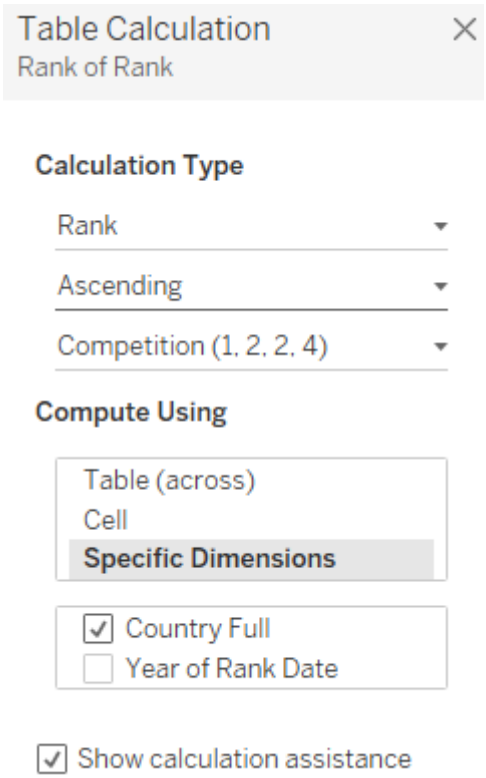
2. Add **`Rank Date`** and **`Rank`** to the **`Columns`** and **`Rows`** shelves. Set **`Rank Date`** to **`Year`** and **`Rank`** to **`Sum`** , as follows:



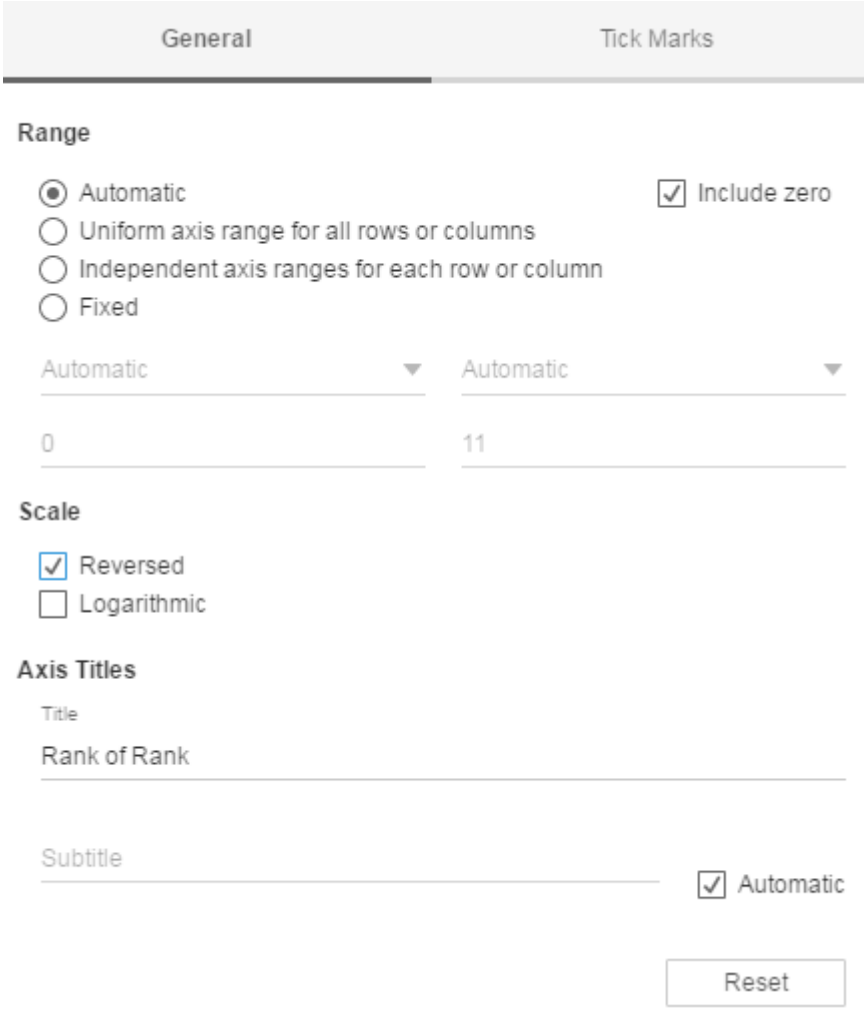
3. Rank in our current dataset is the actual global ranking of each soccer team. Because we want these to show up as 1-10 instead of their country rankings, we apply the ** `Rank Table Calculation` **to the rank field, as follows:



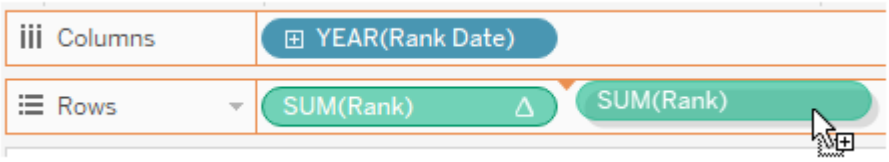
4. Edit the `Table Calculation` to `Rank` in `Ascending` order, and under `Compute Using`, select `Country Full`, as follows:



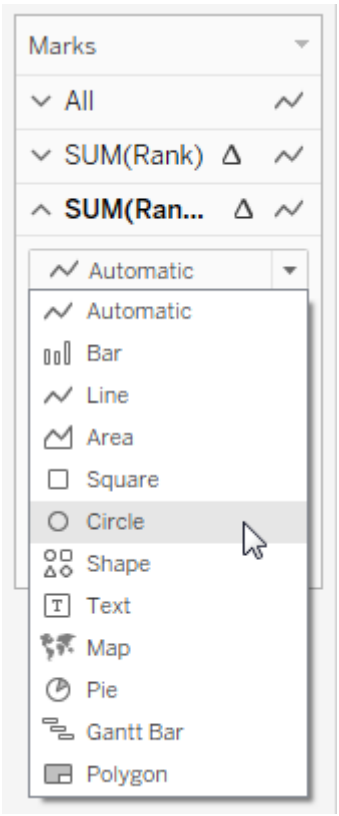
5. Reverse the axis of rank so our top performer is first, as follows:



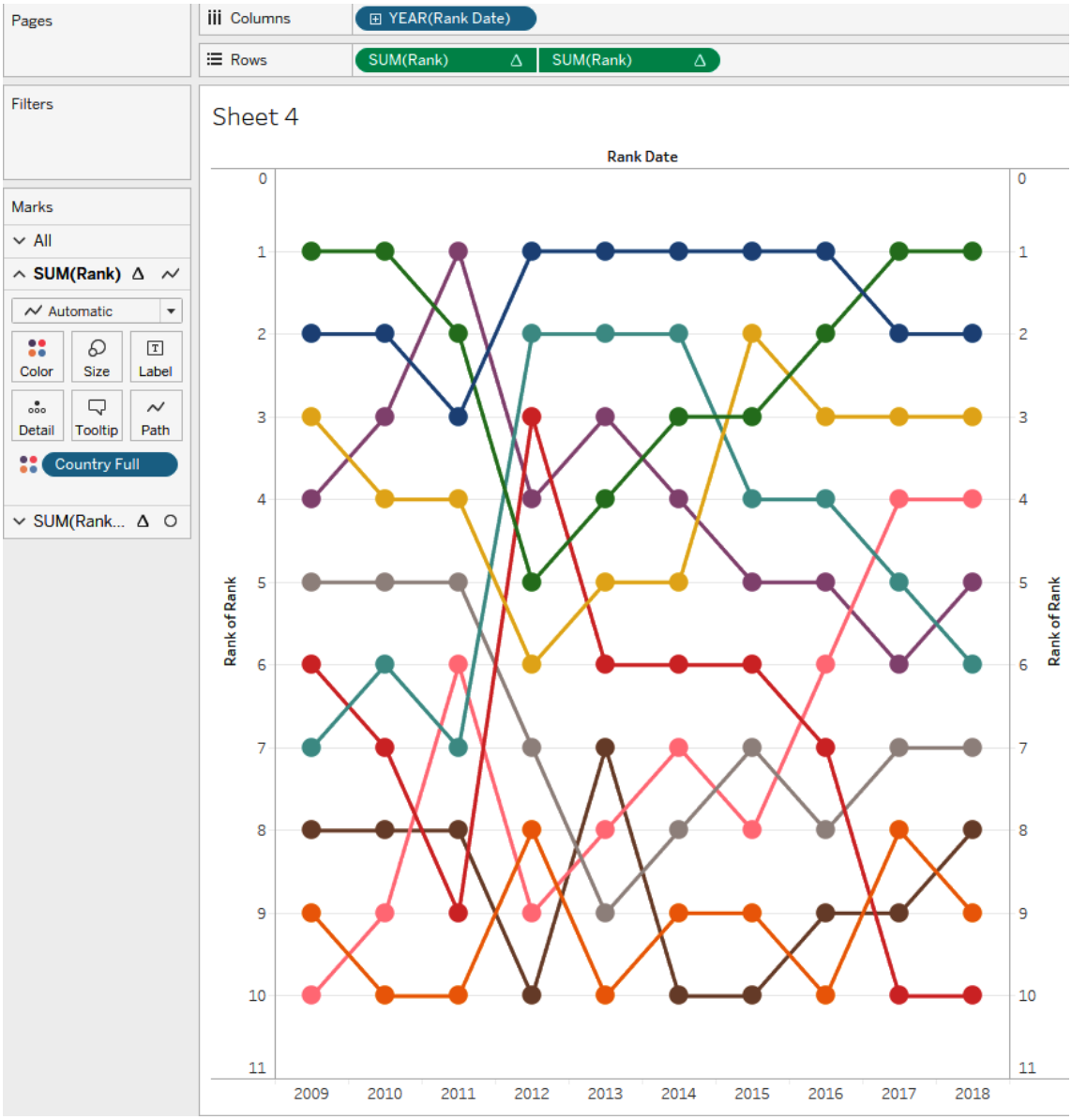
6. Duplicate `Rank` on the `Rows` shelf to create the dot points in our visualization, and make this a dual axis by right-clicking and choosing the option, as follows:



7. Change the default mark for the second rank to `Circle`, as follows:



8. Reverse the axis.
9. We now have a basic bump chart, which we can see in the following screenshot:



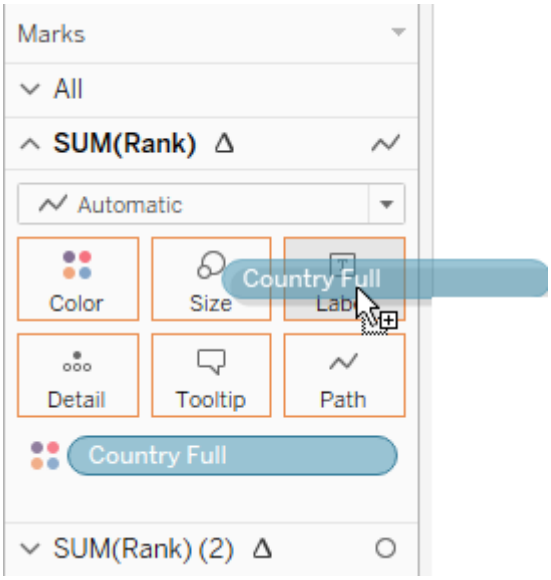
How it works...

In [*step 1*], we put `Country Full` on `Color` in the `Marks` card, which groups and colors the values by `Country`. In [*step 2*], we use a basic line chart for `Rank Date` and `Rank` by putting them on the `Columns` and `Rows` shelves. In [*step 3*] and [*step 4*], the rank needs to be modified so that we see a sequential order rather than the rank in the dataset. In order for first place to show at the the top of the chart, we reversed the axis in [*step 5*]. Next, in [*step 6*] and [*step 7*], we duplicate rank on the `Columns` shelf and change the `Mark` type to `Circle` to emphasize each yearly rank. In [*step 8*], we reverse the axis as in [*step 5*]. In [*step 9*], we merge the line (`Automatic`) and `Circle` marks by using dual axis, and, after some formatting and size changes, we have a basic bump chart.

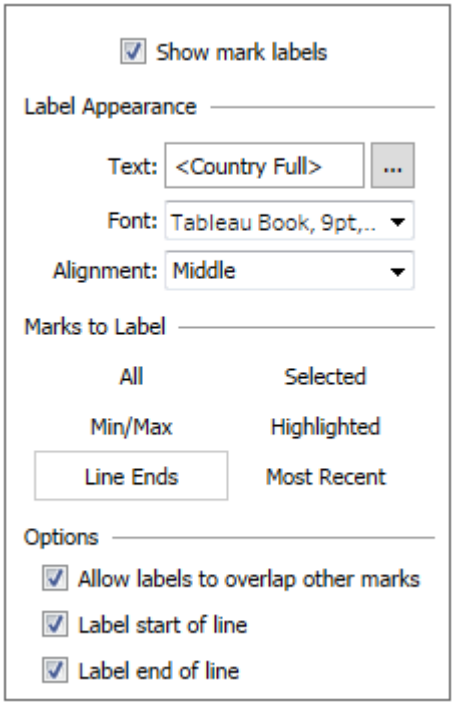
There's more...

We can label the start and end of the lines by taking the following steps.

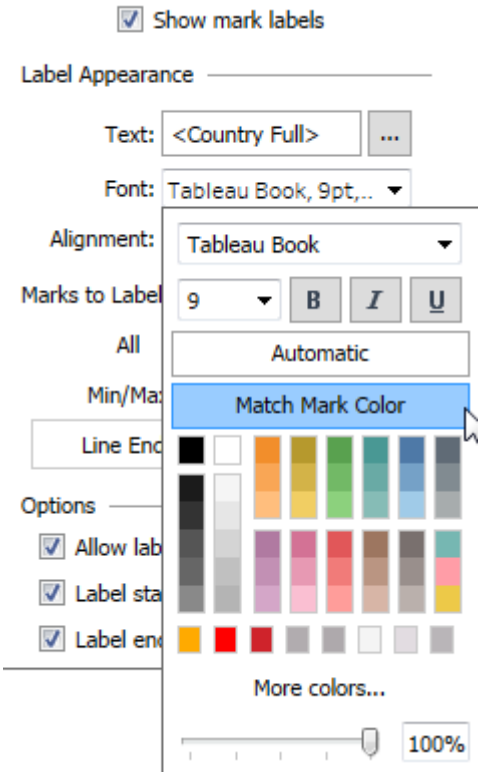
1. In order to easily see where a country starts and ends in the ranking, label the lines. While in the `SUM(Rank) Marks` card, move `Country Full` to `Label`, as follows:



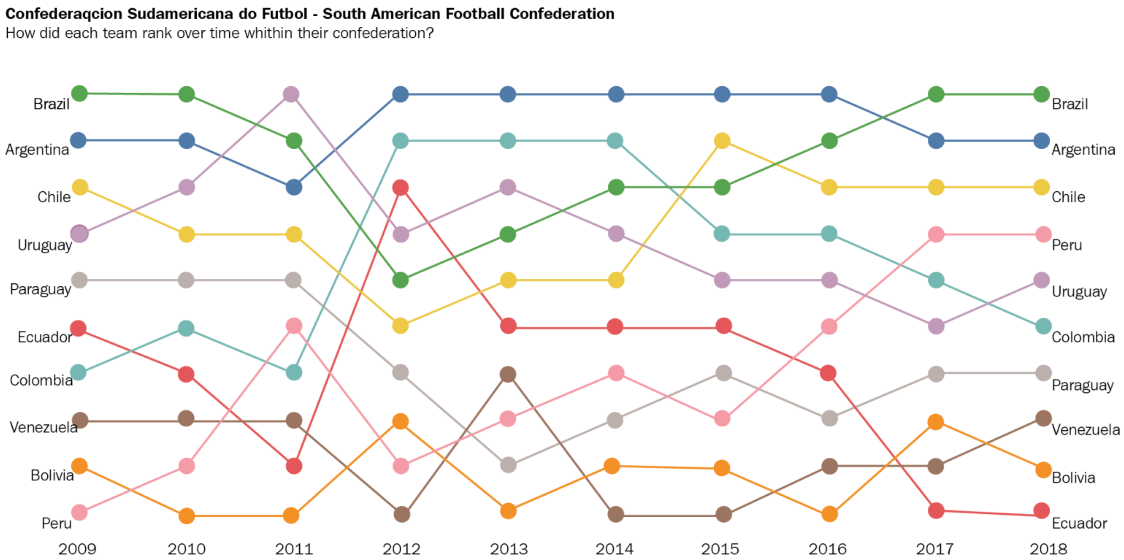
2. Choose `Line Ends`. You may need to change the alignment or font sizes, as follows:



3. Match the text color with the line color, as follows:



4. After the labels are applied and some additional cleaning up, here is our final chart:



Sparklines chart

Sparklines are compact visualizations. They show patterns over time and are essentially line charts without the axis. Edward Tufte explains that:

["*Sparklines are datawords: data-intense, design-simple, and word-sized graphics.*"]

In other words, we recognize their form and the information being relayed as easily as we can words on a page---"word-like graphics".

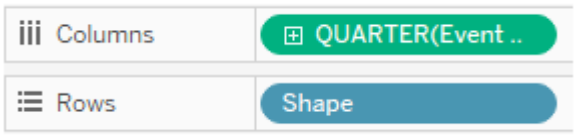
Getting ready

In order to prepare for the Sparkline recipe, we will need a basic understanding of line charts. In this recipe, we rely heavily on the format lines menu and editing the axis.

How to do it..

In this section, we use UFO sighting data to make Sparklines. Open the Sparkline packaged workbook to follow along. To follow along, use `nuforc_events.csv` and `Sparklines.twbx`.

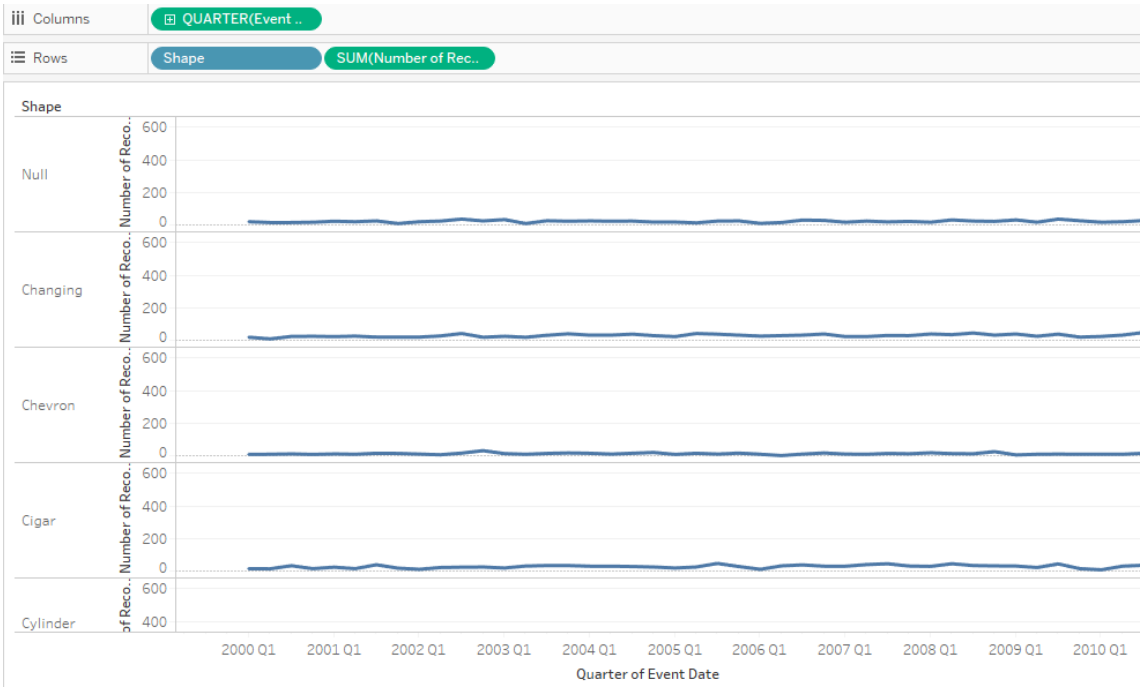
1. To begin, move the event date to **`Columns`** and shape to **`Rows`**. Because we have several years of data, we're filtering the dataset to 2010-2016 and displaying quarters, as follows:



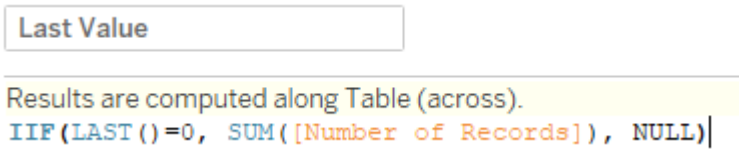
Note

We use the green pill for **`QUARTER(Event Date)`** as opposed to the blue one because, in this scenario, the date is continuous rather than a discrete value. In other words, we want date to appear as a value on an axis rather than a value to be grouped.

2. Move the measure to the **`Rows`** shelf. For this dataset, we are using the number of records, as follows:

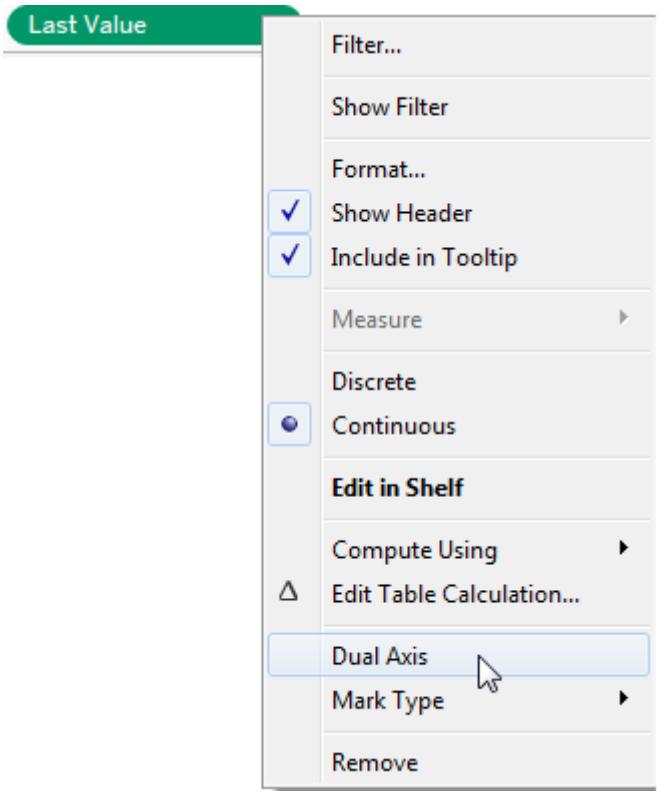


3. Create a measure called **`Last Value`** to indicate the final value in the Sparkline, as follows:

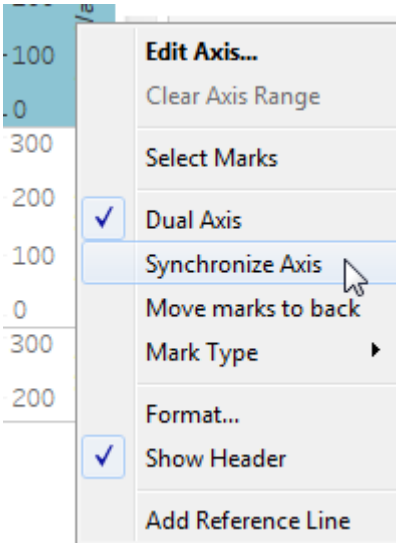


4. Move `Last Value` to the `Rows` shelf, as follows:

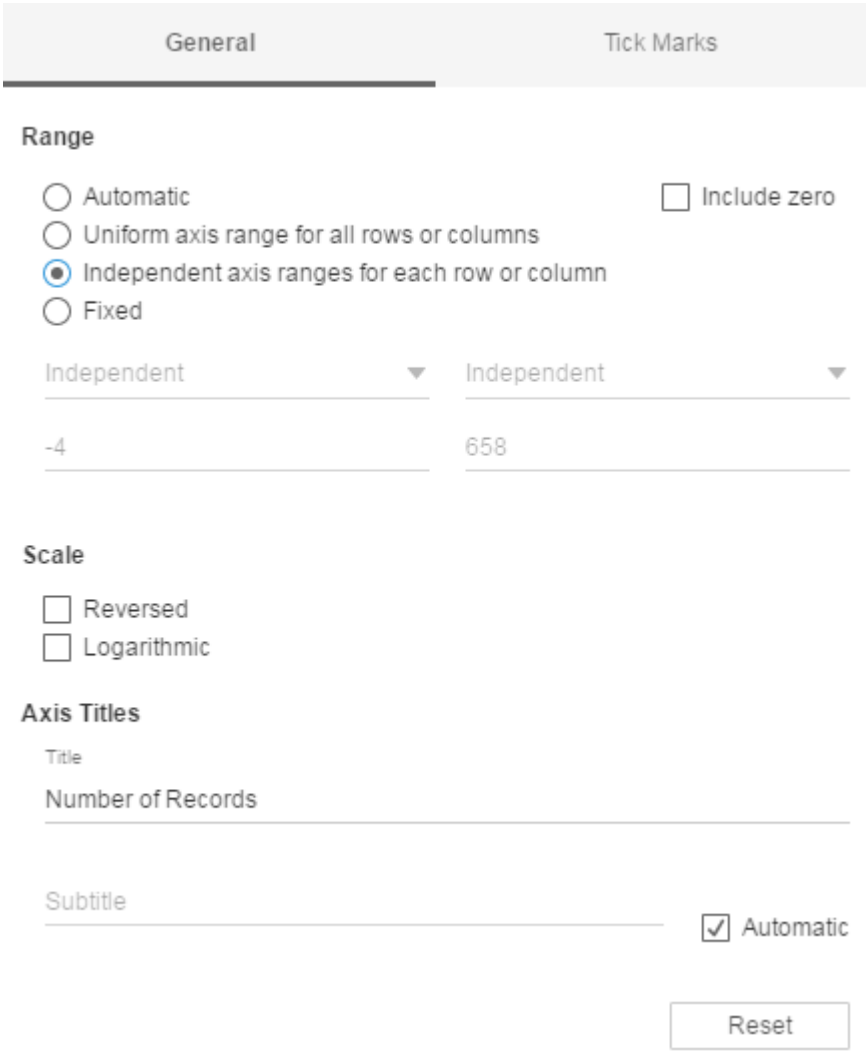
5. Right-click on `Last Value` and select `Dual Axis`, as follows:



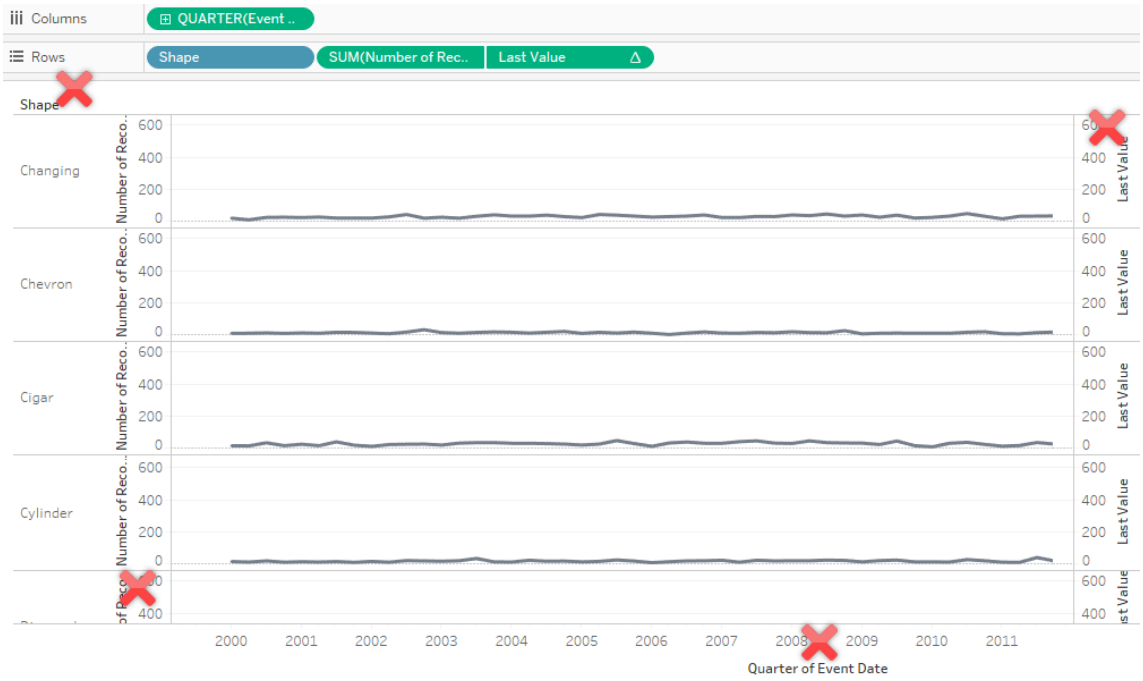
6. Right-click on the axis in the visualization and choose `Synchronize Axis`, as follows:



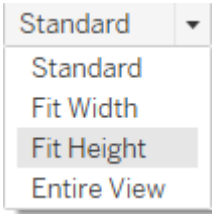
7. Edit the axis by deselecting the `Include Zero` option and choosing the `Independent axis ranges for each row or column` radio button, as follows:



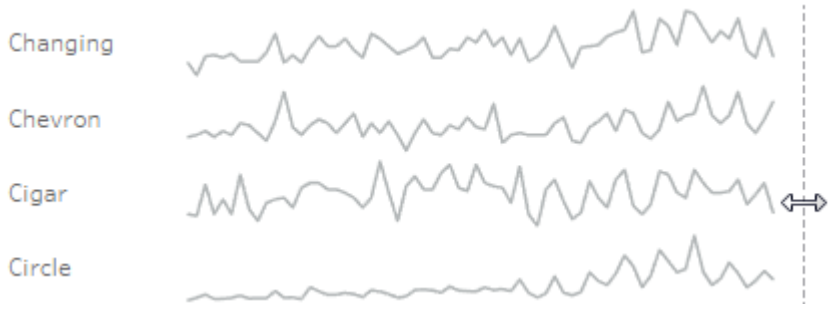
8. Remove all the headers by right-clicking on the visualization where each red x is located and deselecting `Show Header`, as follows:



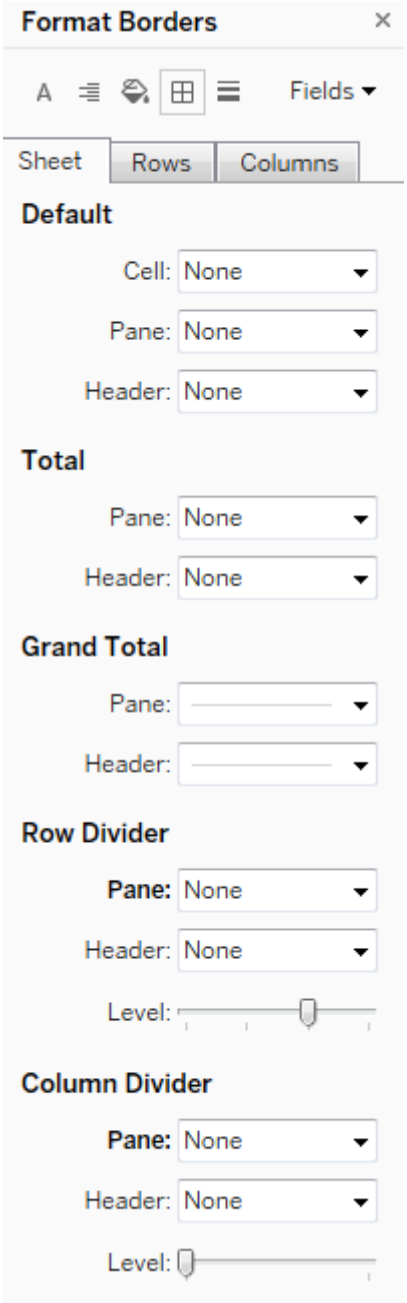
9. Fit the height, as follows:



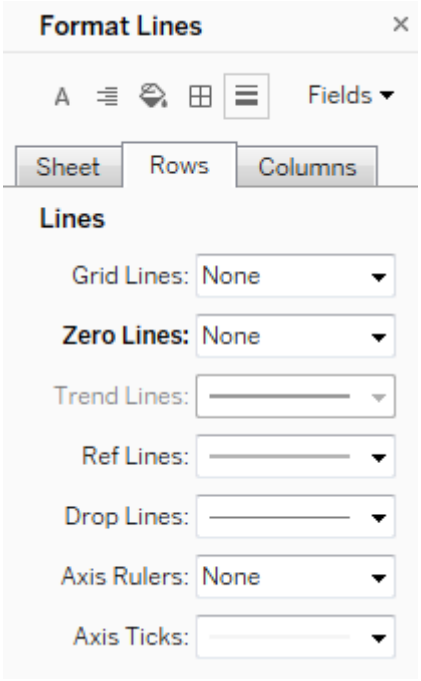
10. Manually resize the width by dragging the column ending inward, as follows:



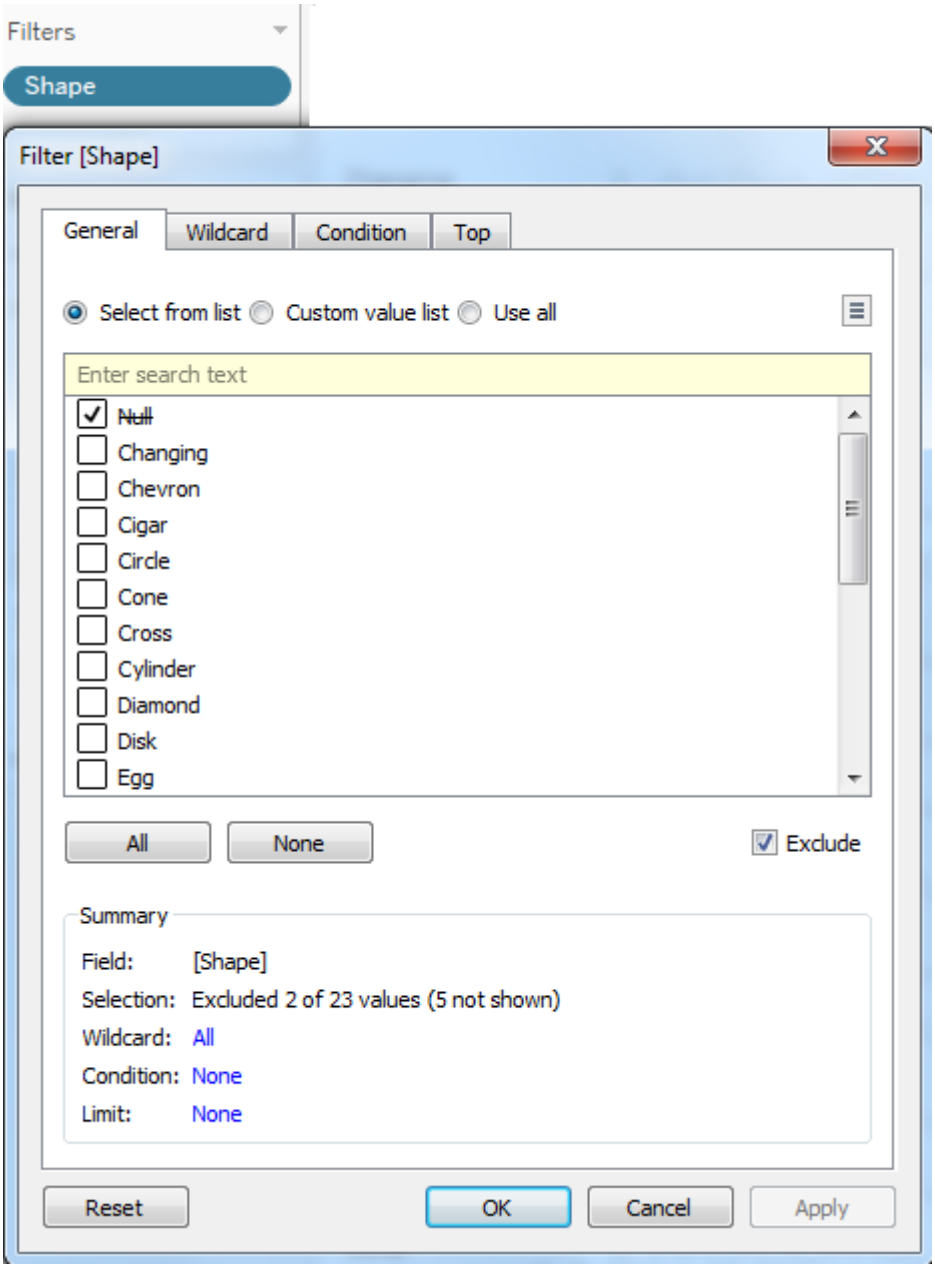
11. Remove all grid lines. From `Format Borders`, choose `Sheet`, and set `Column Divider` and `Row Divider Pane` to `None`, as follows:



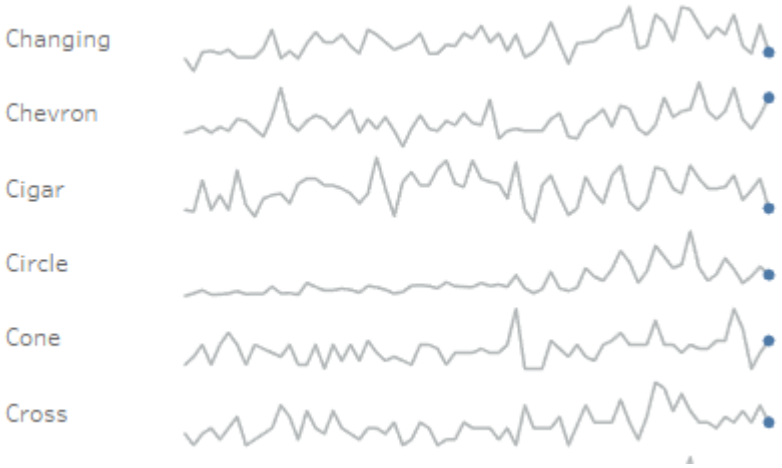
12. From **`Format Lines`**, choose **`Rows`**, and set **`Grid Lines`** and **`Zero Lines`** to **`None`**, as follows:



13. Exclude null shapes by dragging **`Shape`** to the **`Filters`** card and select **`Null`** and **`Exclude`**, as follows:



14. The Sparkline chart should appear as follows:



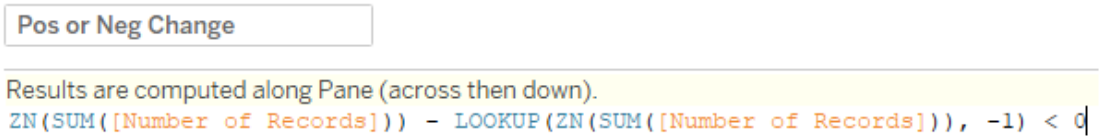
How it works...

In [*step 1*], we put the event date on the `Columns` shelf and shape on the `Rows` shelf to create a long list of line charts. In [*step 2*], we put the number of records on the `Rows` shelf. In order to highlight the final value in the Sparkline, we create a calculation for the final value in [*step 3*], put it on the `Rows` shelf in [*step 4*], and, in [*step 5*], make it a synchronized dual axis chart. Each of the line charts should appear independent so we can see changes for each category relative to itself and not to each other. In [*step 6*], we achieve this by editing the axis---excluding zero and making the axis independent for each row or column. In [*step 7*] to [*step 10*], we remove and hide all the headers, all the grid lines, and shrink the width and height considerably. As a result of the formatting, we see the final Sparklines in [*step 11*].

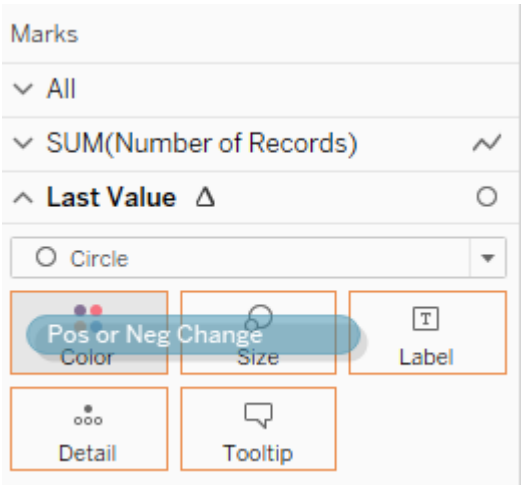
There's more...

Instead of adding a single color at the end of the Sparkline, we could make one that shows whether the last value is higher or lower than the one before, or even display the maximum values, by taking the following steps.

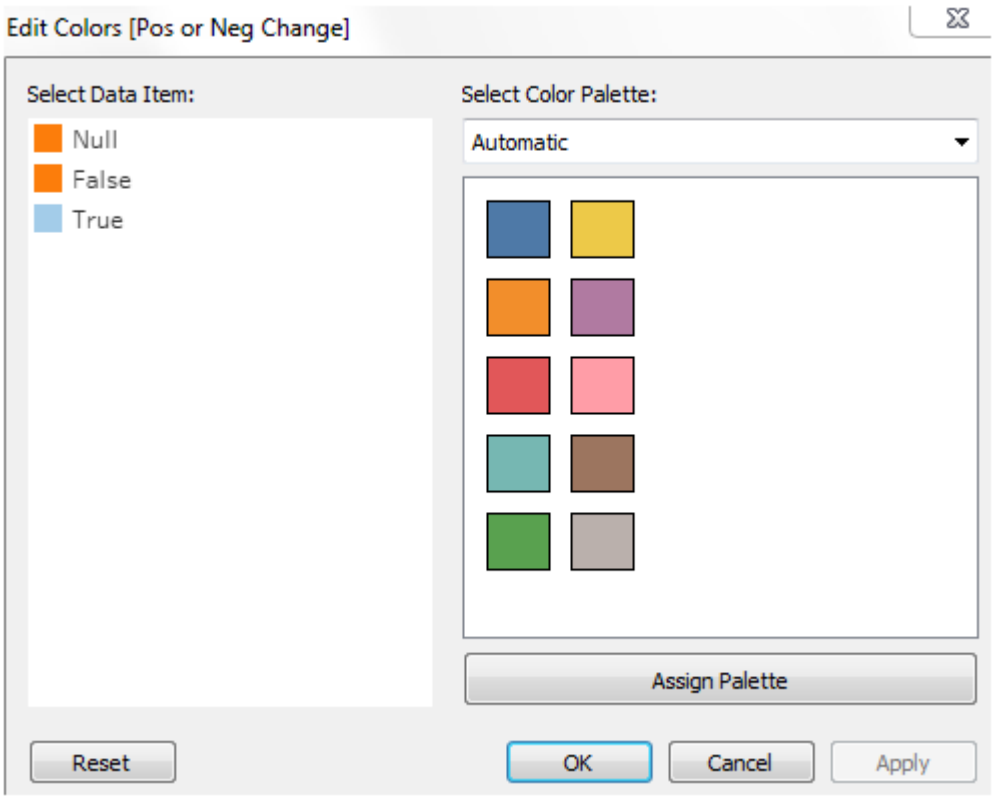
1. In order for the end of line indicator to change color, we need to make the calculation that we can see in the following screenshot:



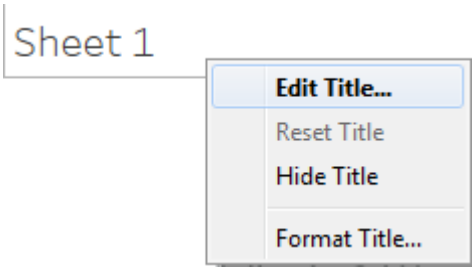
2. Move this calculation to the `Marks` card, as follows:



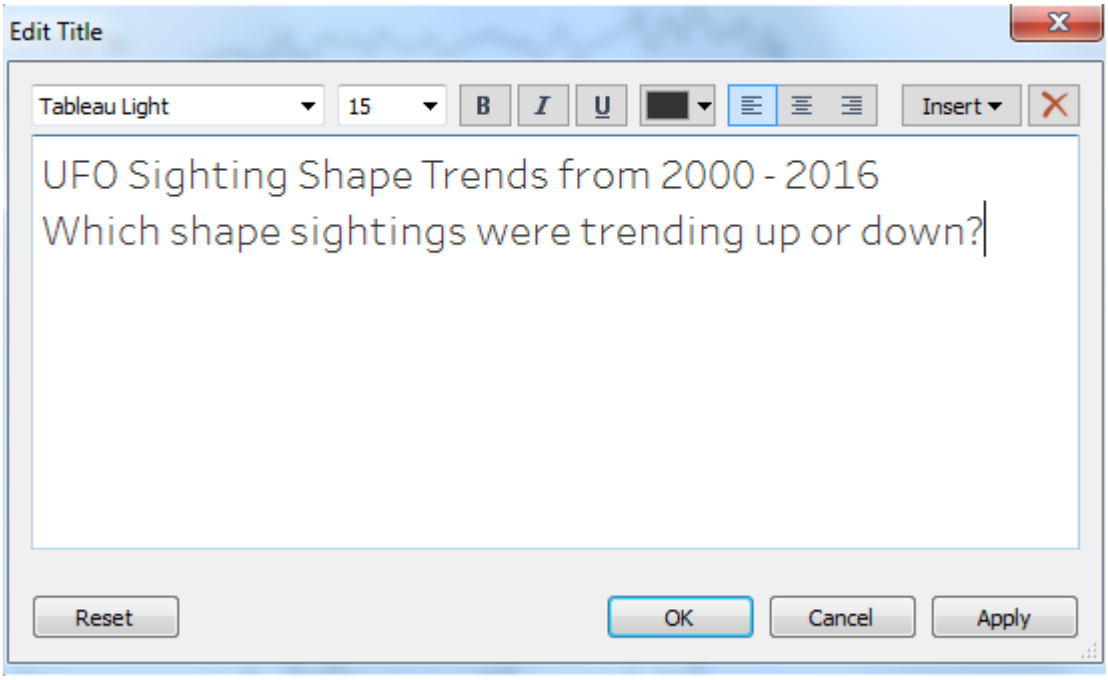
3. Edit the `Color` option, as follows:



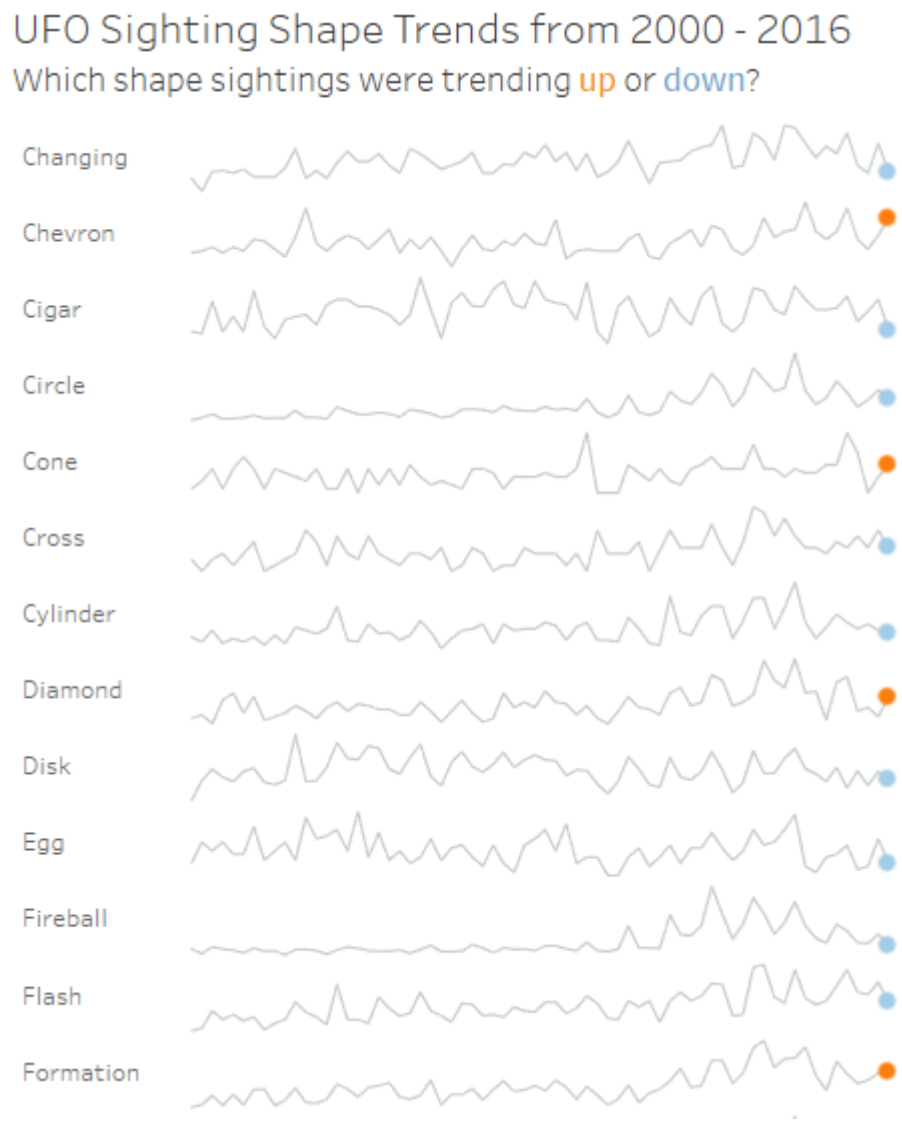
4. Format the title, by right-clicking and choosing `Edit Title`, as follows:



5. Type in the following text, and modify the size and colors as you would in any text editor:

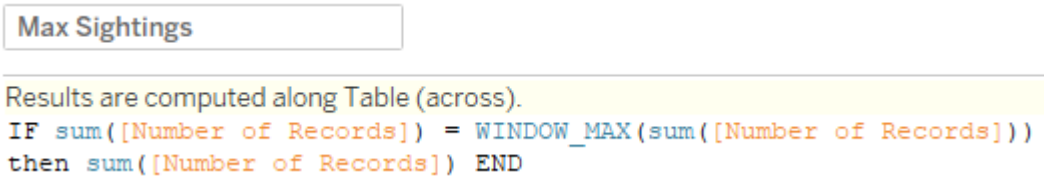


6. Finalize the formatting by further tweaking the sizes, colors, and transparencies, as follows:



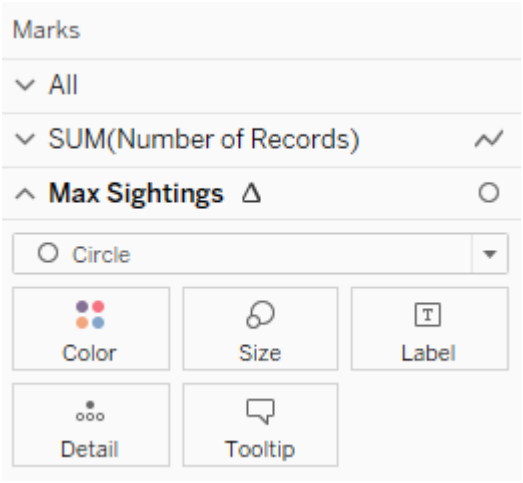
We can also highlight the maximum value instead of the final value.

7. Create the following calculation:

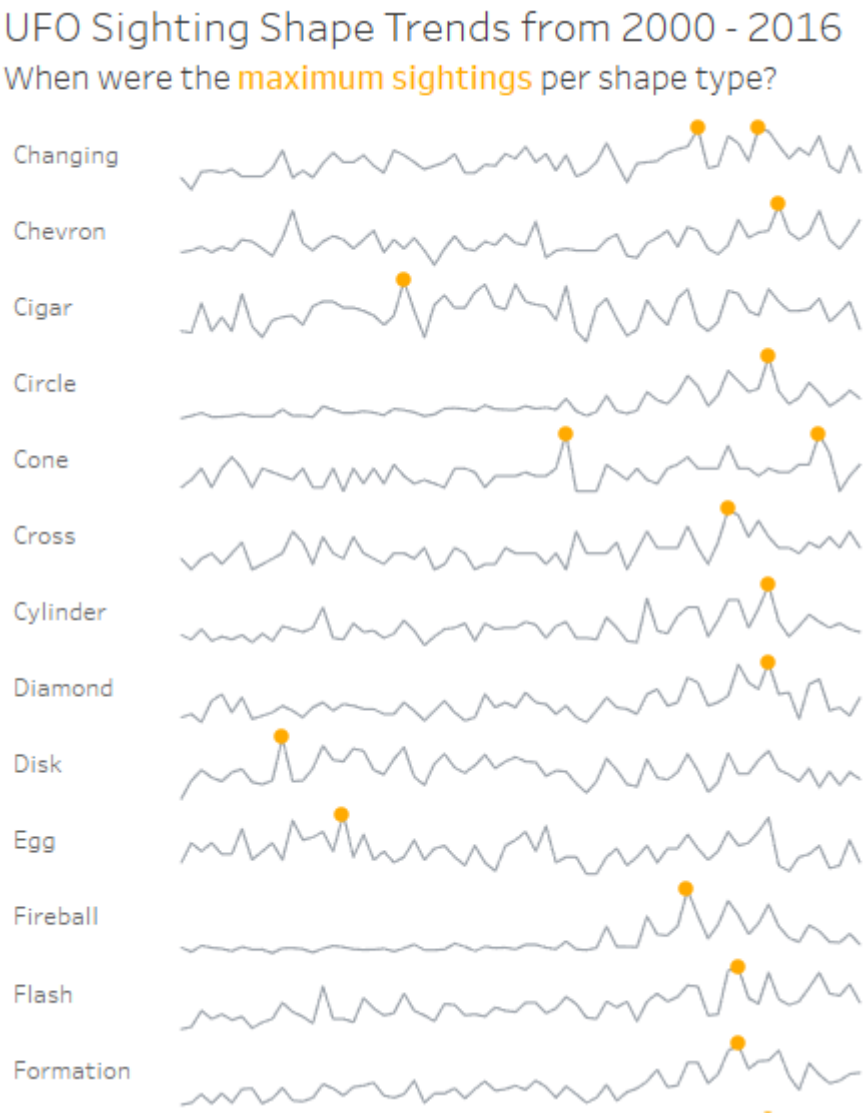


8. Substitute the final value measure with the max sightings one, as follows:

9. Format the color to fit your style, for example:



10. We can see the final Sparkline chart in the following screenshot:



See also

- Read any of Edward Tufte's books. He is said to be the creator of the Sparkline.

Donut chart

Donut charts are pie charts with a hollow center. They are best used when communicating percentages, and can help the distinction of a few categories pop.

Note

Donuts are best used when there are only a few categories to be displayed in the donut. There are other visualization types that would show many categories better, such as horizontal bars to name but one.

Getting ready

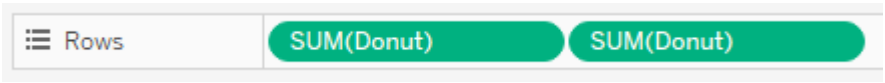
In this recipe, we extend the pie chart with a dual axis to create a donut chart.

How to do it..

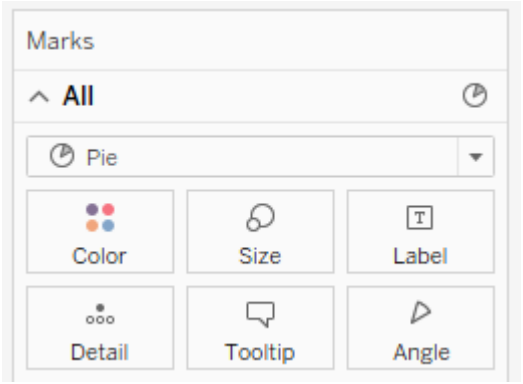
In the next example, we're going to work with avocado data and show which cities eat a higher percentage of organic versus conventional. Use `Donut.twbx` to work through the example.

1. Create a calculation called `Donut` with zero hardcoded as the value, as follows:

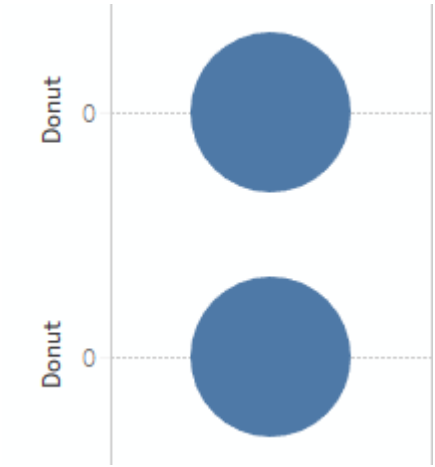
2. Place **`Donut`** on the **`Rows`** shelf twice, as follows:



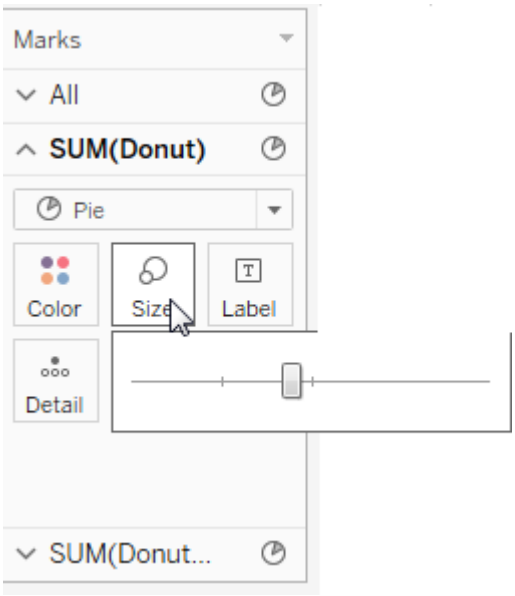
3. In the **`Marks`** card for all values, change the default mark from **`Automatic`** to **`Pie`**, as follows:



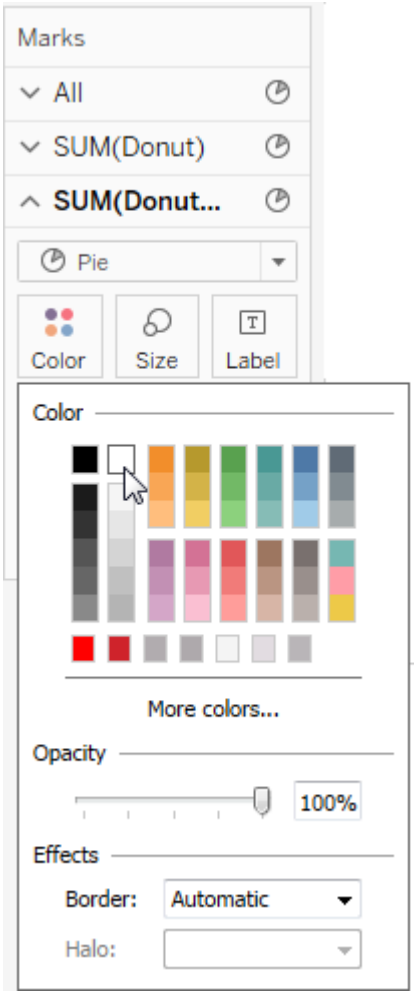
In the following screenshot, we can see how our chart should look:



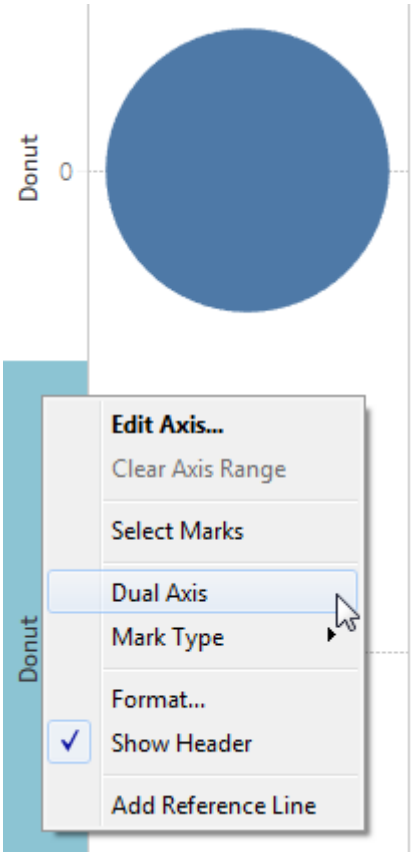
4. Next we will do some basic formatting by making the first donut larger, as follows:



5. Change the color of the second donut to the background color, as follows:



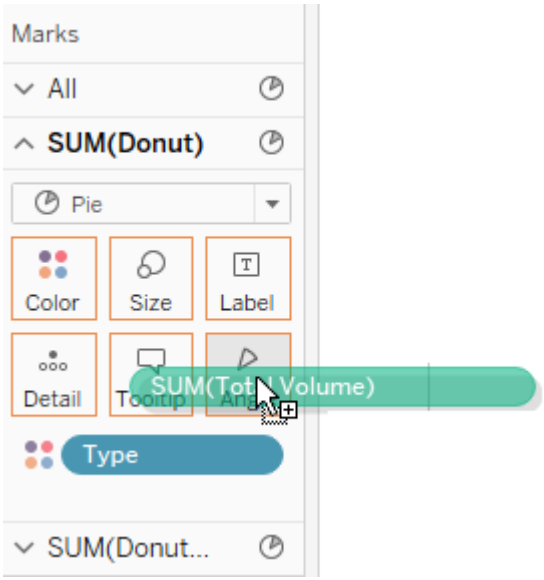
6. Create a dual axis chart by right-clicking on the axis and choosing `Dual Axis`, as follows:



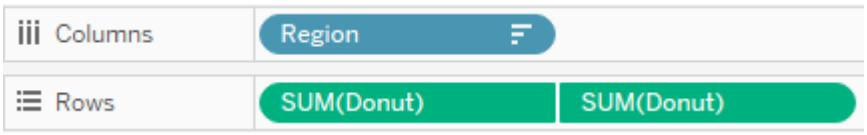
Note

Right-clicking on the second donut pill and choosing `Dual Axis` is an alternative method. At this point in the recipe, we now see a ring, as shown in the following screenshot:

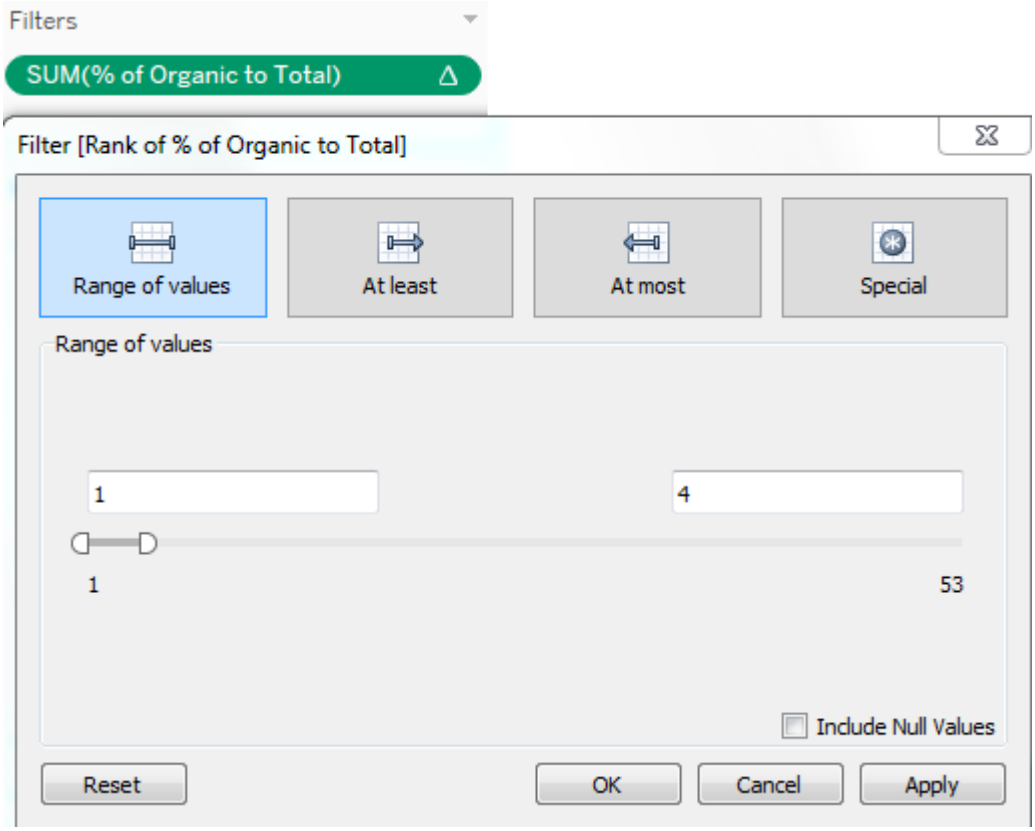
7. In the `Marks` card for the first donut, move the category you want to display as slices of your pie chart to color. In our example, that's type---organic or conventional. Because we haven't defined the angle, you should see an even split.
8. Next, we'll define the portion of each slice by moving the measure to `Angle`. In our example, this will be `Total Volume`:



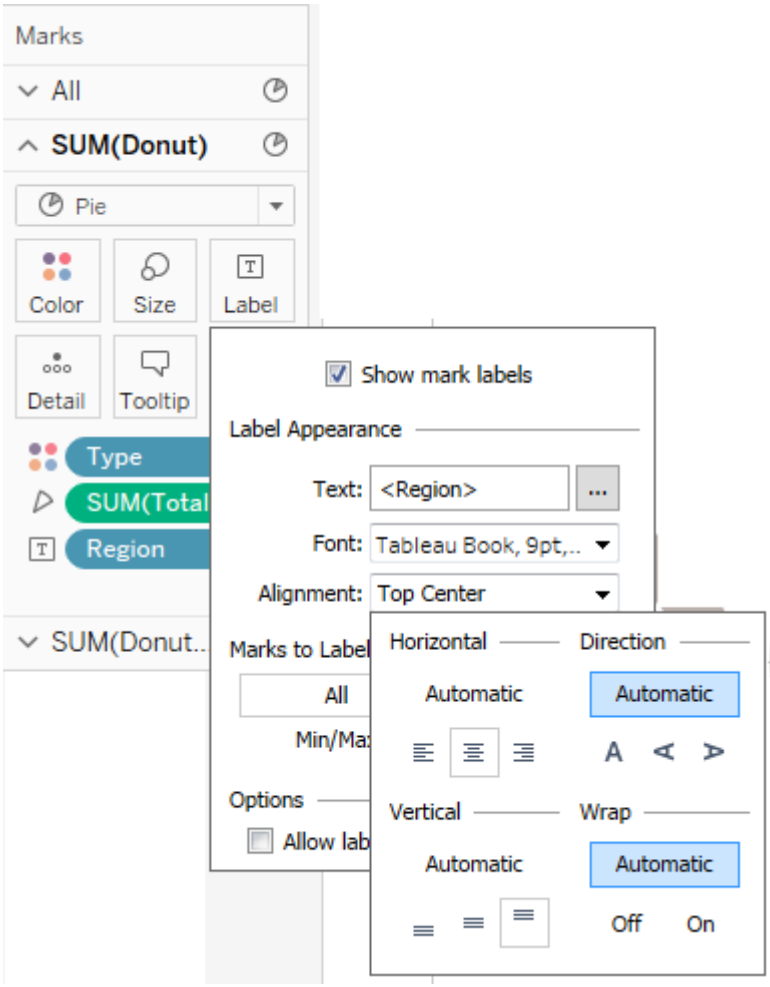
9. To create the horizontal series, put a dimension in the `Rows` . In our example, this would be `Region` :



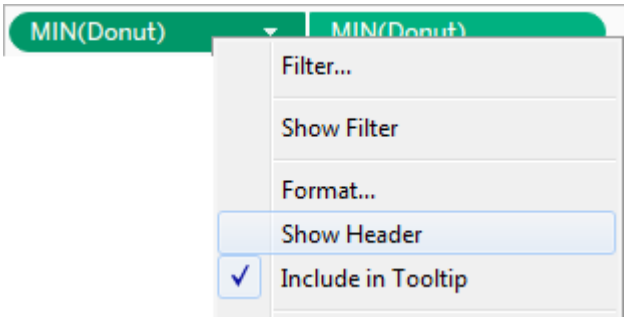
Because we have so many regions, I've also added a top rank filter so we only see a few, as follows:



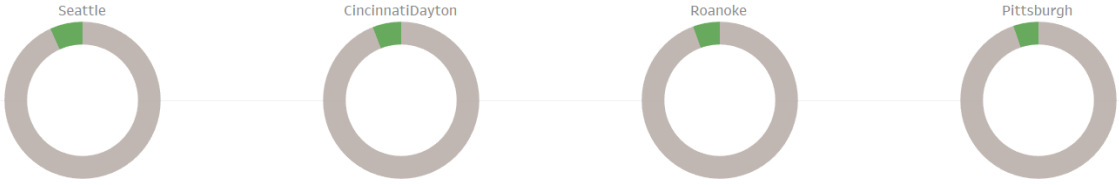
10. In the ** `Marks` **card for the first donut chart, move **`Region`** to **`Label`** and change the alignment to be top and centered, as follows:



11. Remove the axis label, right-click on each donut pill, and deselect `Show``Header`, as follows:



We can see our donut chart in the following screenshot:



Note

Donuts can be effective and fun when comparing two or three categories. Shy away from this chart type when comparing more.

How it works...

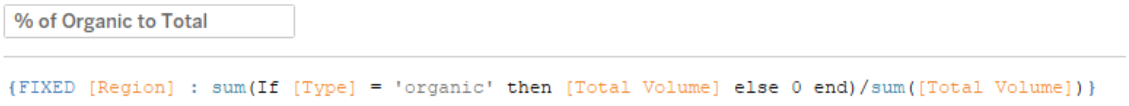
In [*step 1*], we create a calculation called donut and place it on the `Rows` shelf twice in [*step 2*]. In [*step 3*], we change the `Marks` card from `Automatic` to `Pie`. In [*step 4*], we see that we have two pies. In [*step 5*], we modify the format for one pie to be larger than the other pie. In [*step 6*], to hollow out the center, the color of the smaller pie is set to match the background. In [*step 7*], we create a dual axis chart that overlays the smaller circle onto the larger circle. In [*step 8*], we should see a donut chart.

Next, in [*step 9*], we create slices of the donut by moving the categorical dimension to `Color` on the `Marks` card (first donut measure). The angle shows how much each slice takes of the whole; in [*step 10*], we put `Total Volume` on `Angle`. In [*step 11*], we create a series of donuts by placing region on the `Columns` shelf. In [*step 12*], we create titles for each donut by moving region onto `Label` within the `Marks` card. We format the text so its alignment is top and center. In [*step 12*], we have our final donut chart.

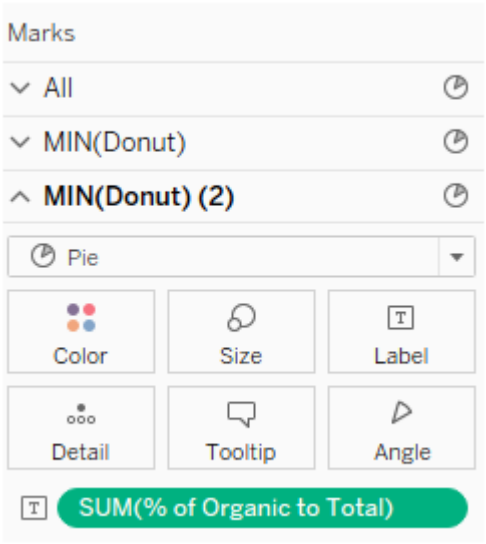
There's more...

The hole of the donut can be used to display a value like the percent of the total.

1. Create a % of total metric, as follows:



2. Place the % of `Organic to Total` on to `Label` in the second donut's `Marks` card. Format `Alignment` to be `Middle Center` and modify the font size and color. Format the decimal places as follows:



The final chart can be seen in the following screenshot:

Motion chart

This section highlights how to use Tableau to page through the data as a movie, frame by frame. Animated charts are useful in showing individual elements, as well as the big picture, in an immersive fashion.

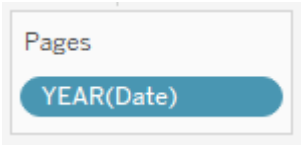
Getting ready

In the motion chart recipe, we learn how to use **`Pages`**.

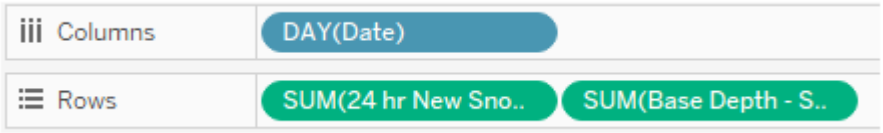
How to do it..

In this recipe, we're playing around with snow data for Snowbird, a ski resort in Utah. Open packaged workbook ski fall animation to follow along, `ski fall animation.twbx` and `Snowbird - Utah.csv`:

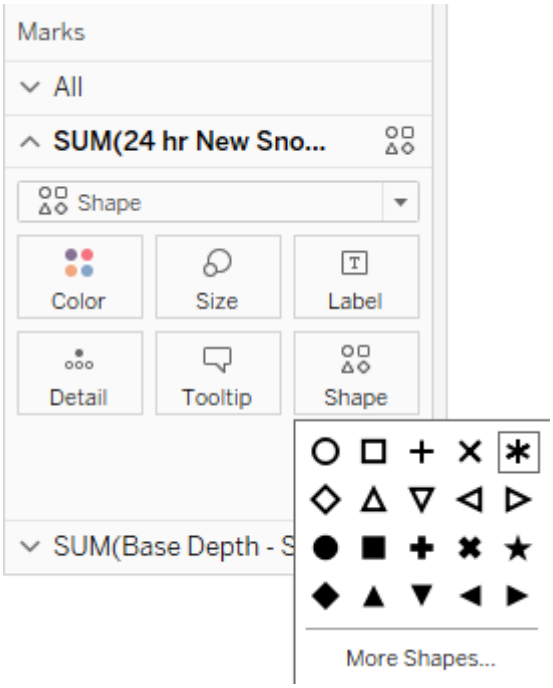
1. Drag a dimension with multiple values to the **`Pages`** shelf. In this example, we use **`YEAR(Date)`**. It works best when you have several values:



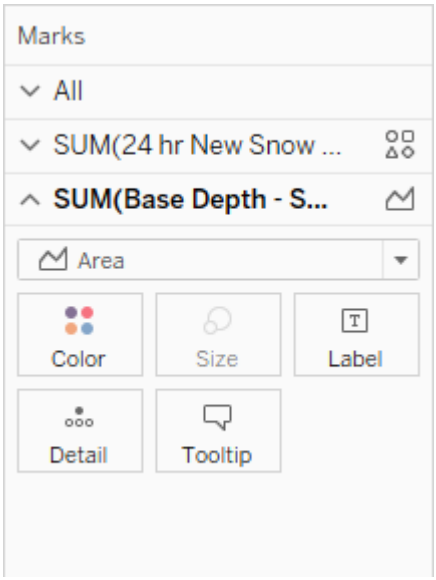
2. Move the date to **`C``olumns`** as **`DAY(Date``)`** and **`24 hour Snow`** and **`Base Depth`** to **`Rows`**, as follows:



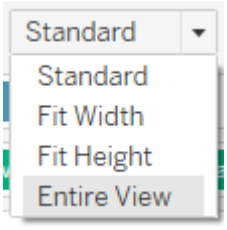
3. Within the **`24 hr New Snow- Split 1``Marks`** card, select **`Shape`** instead of **`Automatic`** and choose the star, as follows:



4. Within the **`Base Depth - Split 1 Marks`** card, select **`Area`**, as follows:

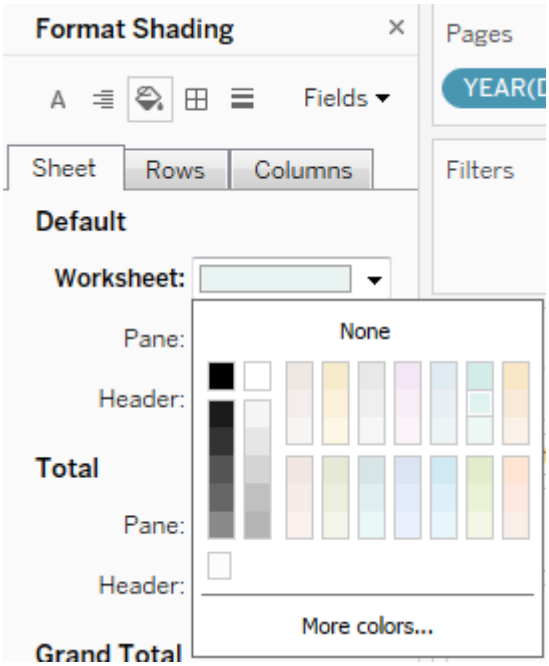


5. Change **`Fit`** to **`Entire View`**, as follows:

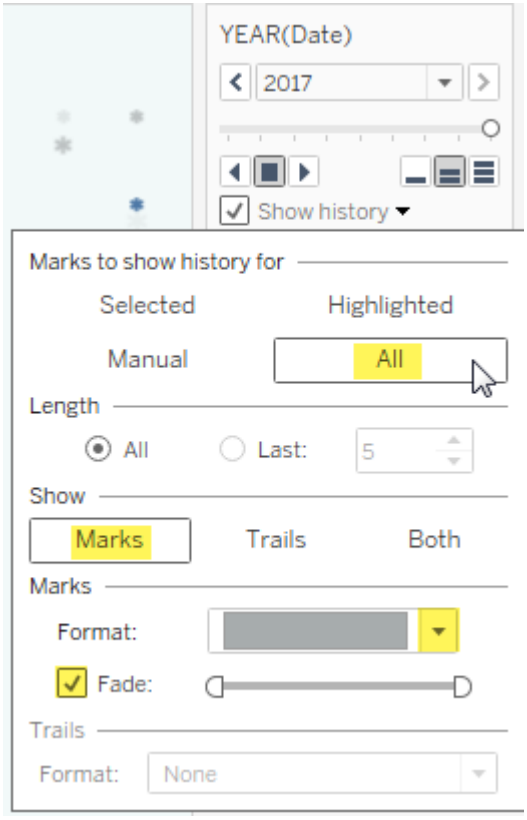


6. Remove all axis headers and grid lines.

7. Change the background to light blue, as follows:



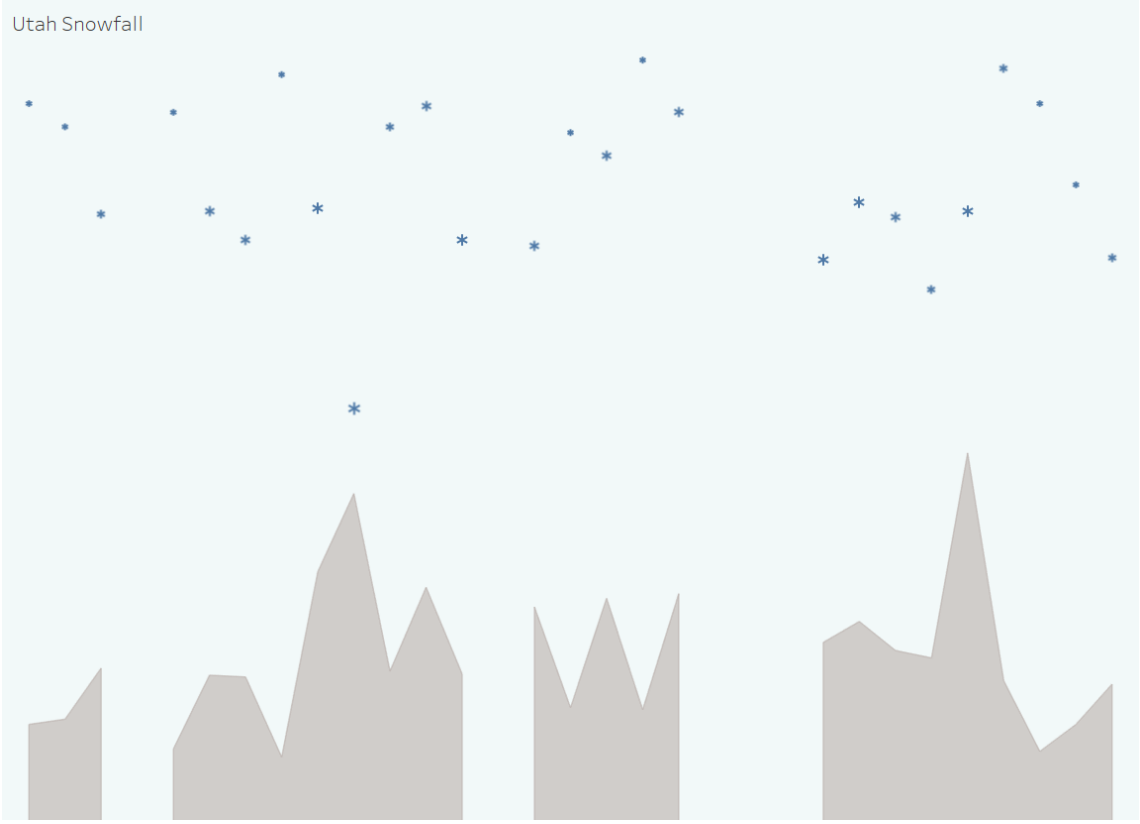
8. Select `All` and `Marks` to show all data points. Select the color grey and `Fade`, to create visual layers, as follows:



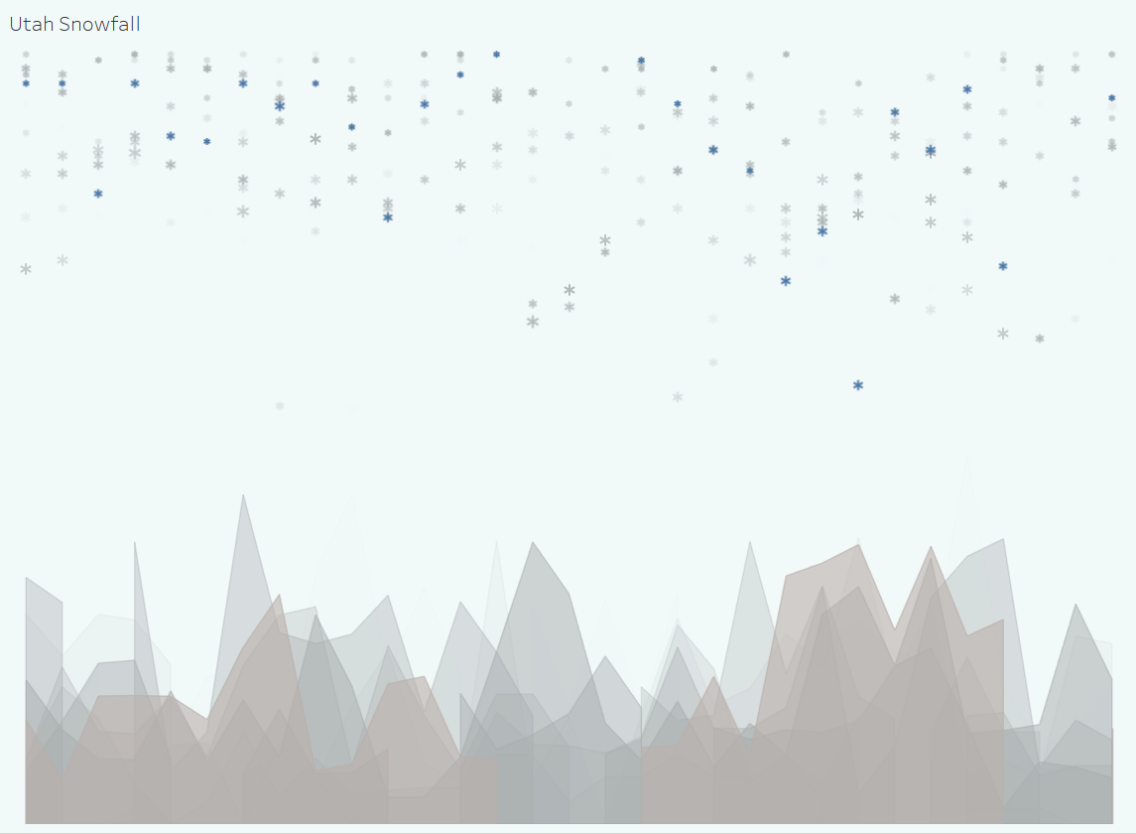
Note

Make sure to continue using split axis, not dual axis. The lines control the speed. The play and stop buttons are intuitive. The part that makes this visualization fun is the history.

9. Here's the first page after we hit play:



10. This is the last page after we hit play:



How it works...

In [*step 1*], we use a time element in the `Pages` shelf to make an animated flip book. The history, illustrated in [*step 3*], allows us to show changes in time or position relative to the previous data point. Because we chose to see all marks with fading, we're able to see a layered ghostly effect.

There's more...

In order to animate automatically when it is published to the server, you must make an animated GIF. You may check the tutorials for your favorite photo editing tool. Here is an example using GIMP, https://www.digitalcitizen.life/how-create-animated-gif-using-your-own-pictures-gimp. Also, play around with trail marks, which will give different effects.

See also

- Please watch any of Hans Rosling's [*Ted Talks*] videos and check out the Gapminder blogs. These videos are famous and fun. Hans Rosling is credited with being the creator of the motion chart.