

## Groups

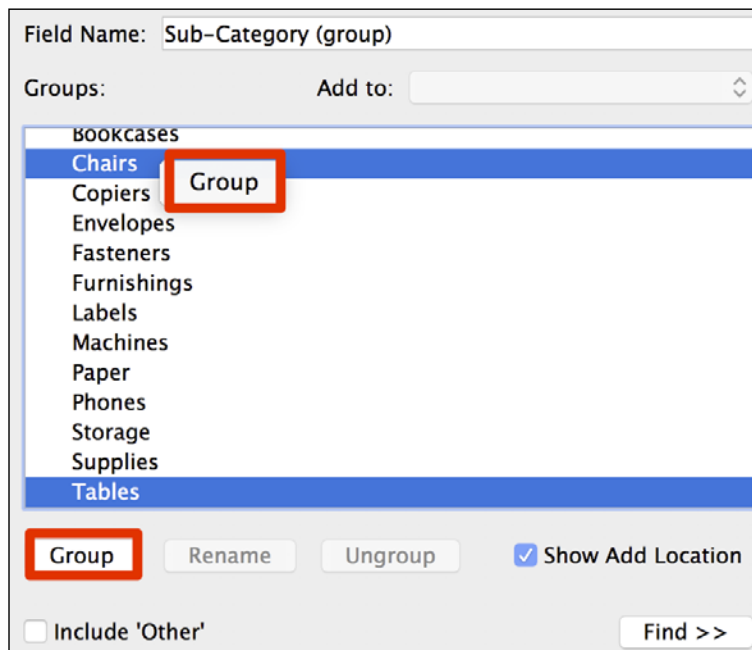
A group is a way to create a new dimension that gathers different values of another dimension. Additionally, a group is static; this means that you need to modify it manually.

A group is characterized by the  icon.

There are two ways to create a group; first, **manually**, with a menu – this is for when you know in advance how to gather the values. Second, you can create a group **visually** in the View.



Let's start by creating a group, manually, with Sample - Superstore:

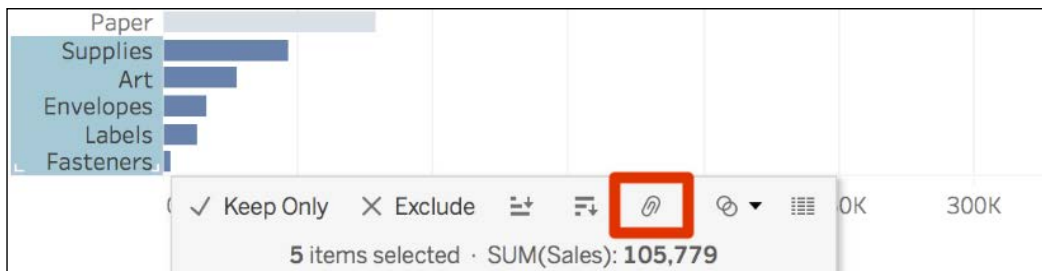
1. Right-click on **Sub-Category**, go to **Create**, and then click on **Group**. A new window opens where you see the list of the values in that Dimension. In this menu, you can manage the groups.
2. To create a new group of values, select some values and click on the **Group** button at the bottom (or with a right-click). Then, create a group with **Tables** and **Chairs**, as demonstrated in the following screenshot:



3. By default, Tableau names the group with a concatenation of the name of the values. However, it's possible to rename it by clicking on the **Rename** button. Rename the group Desk Furniture.
4. After clicking on **OK**, a new Dimension is created. If you didn't change the name, then it is called **Sub-Category (group)**.
5. To test the group, try creating a simple visualization – double-click on **Profit**, and then double-click on **Sub-Category (group)**. You should see your new group, Desk Furnitures, among the list.

If you have too many values and you want to group them, you can use the visual way. Let's demonstrate how to do this with Sample - Superstore:

1. On a new Worksheet, double-click on **Sales**, and then double-click on **Sub-Category** to create a simple visualization.
2. Use the button in the toolbar to swap (  ) and sort (  ) the values as descending. As you can see, there are some sub-categories with small sales. They are not important, so we will want to group them together.
3. Select the five bottom values by clicking on their names in the header (note that it is crucial that you do not select the bars). Then, in the menu that appears, click on the group icon to create a group. The icon is highlighted in the following screenshot:



4. Again, you can rename the group by editing it on the **Data** pane, or by right-clicking on it in the View to edit the aliases. Choose your favorite way and rename the group Small Sales.

Creating a group visually is an excellent way to reduce the number of insignificant values and to help you to focus on what matters. Another perfect use case is when there is a typo, and you want to combine the wrong value with the correct one.

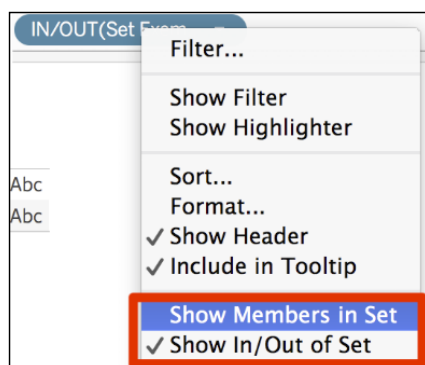
As we said at the beginning, a group is not dynamic. However, sets are dynamic, so let's demonstrate how to use them.

## Sets

Sets are a Tableau element. A set is created from a Dimension. Unlike groups, sets are dynamic. With sets, the values are either **In** the set or **Out**.

Sets have a special icon: .

Sets have two ways of being displayed in a Worksheet. By default, a set will return **In** or **Out**, but if you right-click on its pill, then you can see that it's possible to show the members that are in the set:



---

# An Introduction to Parameters

As you make your way through *Practical Tableau*, and especially **Part III**, there is one functionality that you will hear about over and over again: parameters. Parameters are user-generated values that are not attached to a dataset, and due to their flexibility, are the solution to several of the handiest tricks in Tableau. Other than filters, parameters are also one of the most powerful tactics in Tableau for transferring control from the author to the end user. As discussed in **Part IV**, providing this type of user interaction not only improves the engagement with your dashboards, but it improves the retention of insights and improves the likelihood that they will be shared.

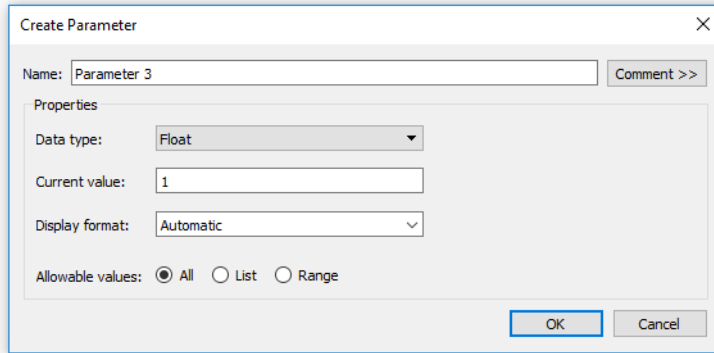
This chapter will use a simple algebra example to illustrate how to create a parameter, how to allow end users to control the parameter you've created, and how parameters work.

## An Introduction to Parameters in Tableau

For this tutorial, imagine that you want to set up a simple equation for 2 multiplied by  $X$ , where  $X$  is the parameter, and the end user gets to choose the value of  $X$ . Tableau will then display the answer of 2 multiplied by  $X$  based on the parameter selection.

For this introduction, it does not matter what dataset you are following along with, but for consistency with the rest of *Practical Tableau*, I will be building a parameter in the Sample – Superstore dataset. There are several ways to create a new parameter in Tableau, including (a) clicking the down arrow in the top-right corner of the Dimensions area of the Data pane and choosing “Create parameter” or (b) right-clicking in a blank space on the Data pane Shelf and choosing “Create parameter.” Once you have

done this, you will be presented with a dialog box where you can design the parameter:



Instead of “Parameter 3,” named that by default in the Sample – Superstore dataset because there are already two other parameters, I will change the name to “Algebra Parameter.” By default, the data type for a new parameter is Float, which means any number including decimals, but there are five additional data types to choose from:

*Integer*

Any whole number (i.e., no decimals)

*String*

Text

*Boolean*

True or false

*Date*

Date without a timestamp

*Date & Time*

Date with a timestamp

The data type you choose for the parameter depends on your use case. For this 2 multiplied by X example, let’s say that we eventually want the parameter to include only whole numbers, so I will change the data type from Float to Integer.

Within the parameter creation dialog box, you also have the ability to designate the allowable values for the parameter. The entries that you choose at this step will eventually determine what choices you and the end users have for the parameter in the final product. The choices are all or any integer, a specific list of integers, or a range of integers. I will choose Range so that I can set a minimum and maximum number as

well as a step size, or multiple, for the parameter. After setting up the parameter to have a range of 1 to 20 with a step size of 1, my parameter looks like this and is ready to be saved by clicking the OK button:

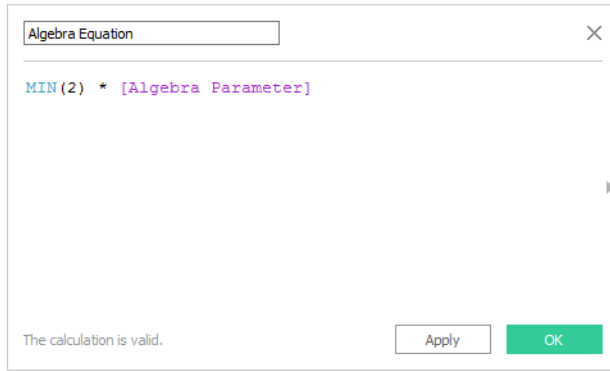
The screenshot shows the 'Create Parameter' dialog box in Tableau. The 'Name' field is set to 'Algebra Parameter'. Under the 'Properties' section, the 'Data type' is 'Integer', the 'Current value' is '1', and the 'Display format' is 'Automatic'. The 'Allowable values' section has three radio buttons: 'All', 'List', and 'Range' (which is selected). Below this, the 'Range of values' section has three checked items: 'Minimum' (value 1), 'Maximum' (value 20), and 'Step size' (value 1). To the right of these fields are two buttons: 'Set from Parameter' and 'Set from Field'. At the bottom right are 'OK' and 'Cancel' buttons.

In order for a parameter to be useful, you must provide Tableau instructions for how to use it. This is accomplished through **calculated fields**. Being that the equation for our use case is 2 multiplied by  $X$ , or our algebra parameter, your first instinct may be to create a calculated field that says:

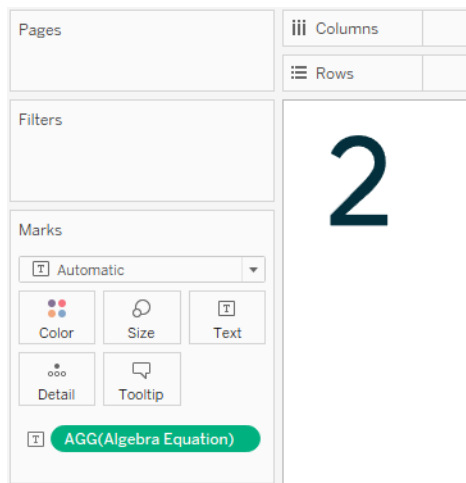
$2 * [\text{Algebra Parameter}]$

Being that the current value of the parameter is 1, we would expect the answer to be  $2 * 1 = 2$ . However, this calculation will be computed for every row in the data, which would not provide the answer we are looking for. Should we add this calculated field to the view, we would get an answer of 19,988, which is 2 multiplied by 1, multiplied by the number of records in the dataset (9,994 for the Sample – Superstore dataset).

To alleviate this, add an aggregation of MIN to the 2 so the Algebra Parameter is guaranteed to be multiplied by 2:



Now if I add my newly created Algebra Equation calculated field to the Text Marks Card to display the answer on a view, I see the answer I am looking for, MIN(2) multiplied by 1 equals 2:



Finally, to provide the power of selecting the value of  $X$  in our example to the end user, you must provide access to the parameter. To do this, right-click the newly created Algebra Parameter and choose “Show parameter control.” A new selector will appear in the upper-right corner of the view that allows the user to choose from the allowable values that were set up when the parameter was created. Notice that the current value is 1, the maximum value is 20, and the numbers can be selected in multiples of 1, just like we set up. Here’s what my final view looks like when I’ve shown the parameter control, moved the parameter control from its default location to below the Marks Cards, and changed the selection from 1 to 7:



In this example, the end user has chosen the number 7, which powered the underlying calculation of  $2 \times 7$ , to produce the answer of 14. Imagine all of the power that can be put into the hands of your end users through the use of parameters!

For just a few other possibilities, see the following chapters:

- Chapter 69, “How to Compare Two Date Ranges on One Axis”
- Chapter 66, “How to Change Date Aggregation Using Parameters”
- Chapter 48, “How to Make a What-If Analysis Using Parameters”
- Chapter 53, “How to Create and Compare”
- Chapter 64, “Allow Users to Choose Measures and Dimensions”



---

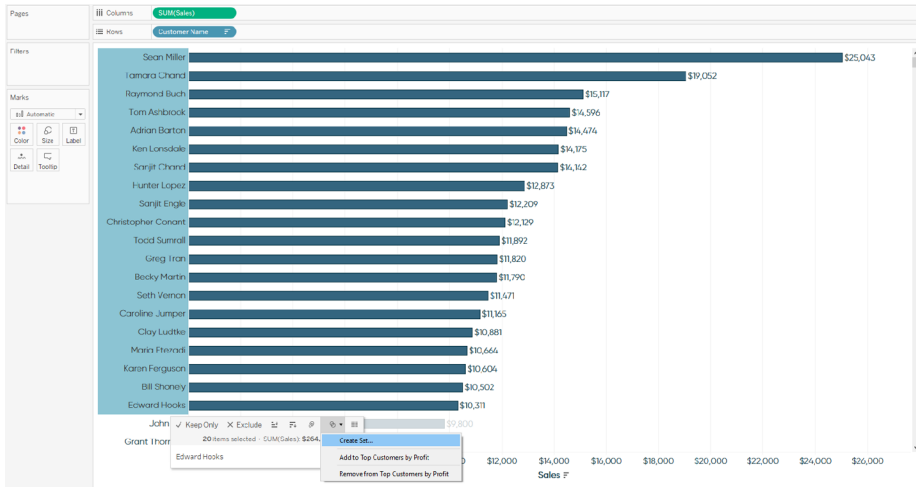
# An Introduction to Sets

Tableau sets allow you to isolate specific segments of a dimension, which can then be used in several different ways to find insights in your data. This chapter provides instructions on how to build sets as well as five different ways they can be used to enhance your analyses.

Sets can be thought of as custom segments, but unlike dimension fields, they are always binary. In other words, you are either in the set or not. Other than that one restriction, sets can be created for just about anything. You can pick individual dimension members to place in a set, have sets be based on quantitative thresholds, created with the top or bottom performing dimension members, and more.

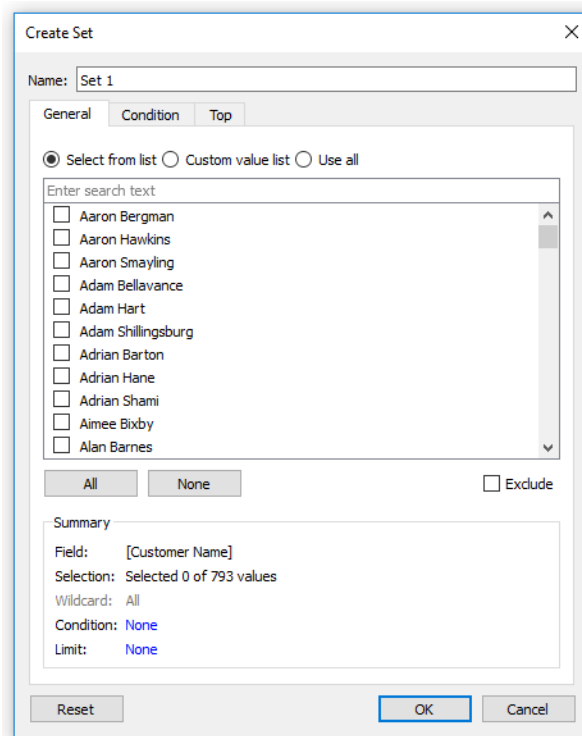
## How to Create a Set in Tableau

Similar to creating filters in Tableau, sets can be created in a couple of different ways. The first and most straightforward method is to simply select the dimension members on a view, then hover over one of the dimension members, click the Venn diagram icon that appears, and choose Create Set. Here's what that would look like if I wanted to create a set from the top 20 customers by sales in the Sample – Superstore dataset:



After clicking Create Set and giving the set a name, the set will appear in its own area on the Data pane. This set that tells us whether each customer name is in or out of the top 20 customers by sales is now available to use in our analyses. Note that sets created with this method are static, so the top 20 will not dynamically change should a new customer enter the top 20.

You can also create a set by right-clicking the dimension the set will be created from on the Dimensions area of the Data pane, hovering over Create, and clicking Set. After following the preceding steps on the Customer Name dimension, the following dialog box appears:

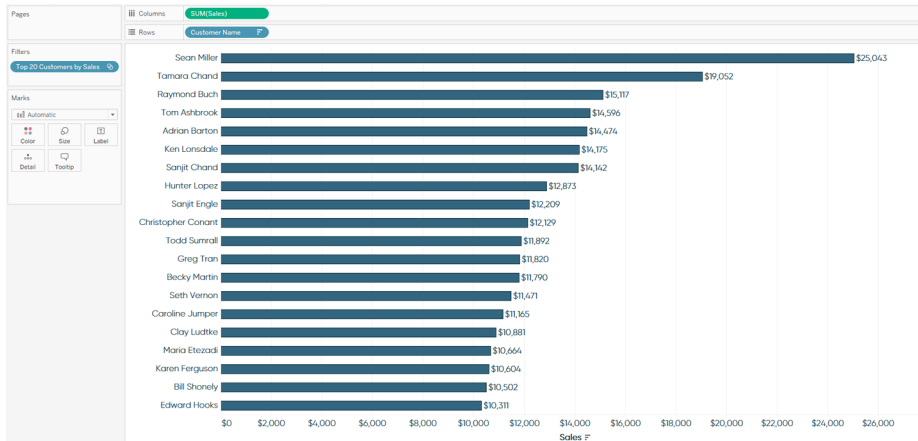


The first tab provides the ability to manually pick and choose the dimension members to be included in the set as we have done with the first method. The remaining two tabs can be used to base the set on a condition or Top N, respectively. Computed sets like this will dynamically change when the dataset is updated. Whichever method you choose to create your set, a new field will be created to use in the following examples.

## Five Ways to Use Tableau Sets

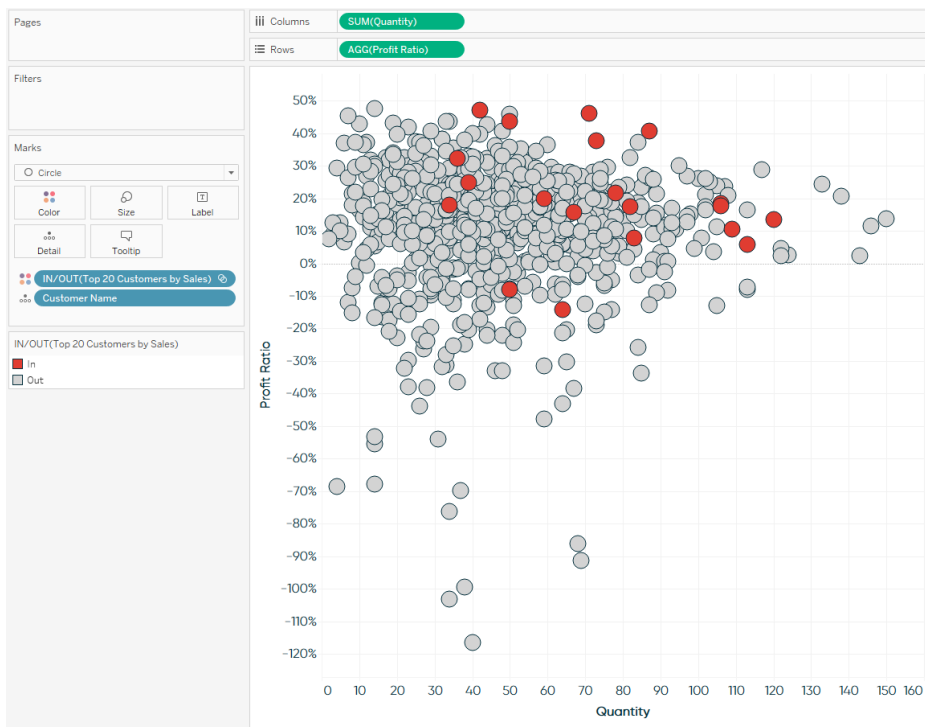
### *As a filter*

Sets can be used as a filter by right-clicking a set from the Sets area of the Data pane and choosing Show Filter. Sets are binary, so when you add them as a filter, you will only have the option to choose whether the marks on the view are in or out of the set. Here's how my bar chart of sales by customer name in descending order looks after showing the filter for Top 20 Customers by Sales and keeping only the names that are in the set:



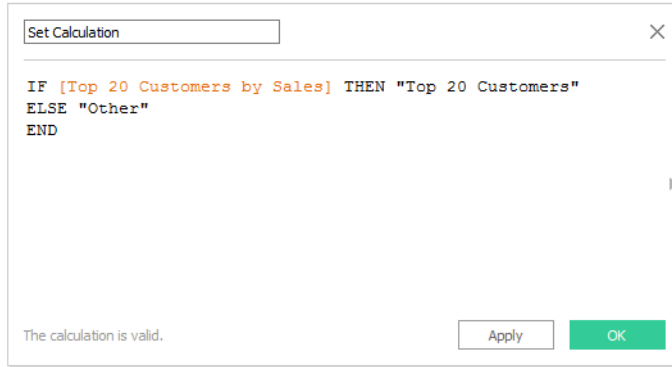
### To encode marks

You can encode marks on a view by a set by dragging the set from the Sets area of the Data pane to the Color Marks Card. Here's an example where I've done this to make the top 20 customers by sales stand out on a scatter plot:



### *In calculated fields*

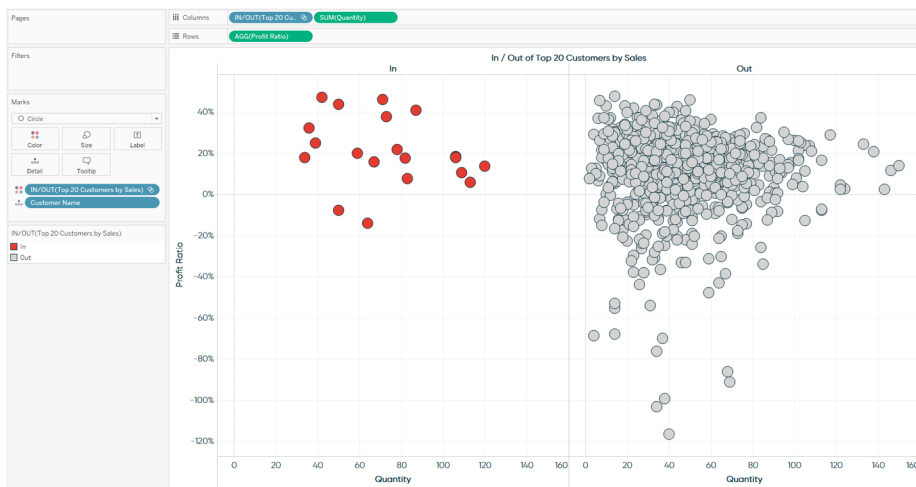
Sets can be used in calculated fields just like dimensions or measures. This way you can treat dimension members differently based on whether or not they are in a set. Here is a simple example being used to rename the sets:



This formula is saying that if a customer name is in the set, name the set “Top 20 Customers”; otherwise call it “Other.” Note that this example is used to show you that sets can be used within calculated fields, but this same renaming could have been accomplished by editing the aliases of the set.

### *As dimension fields*

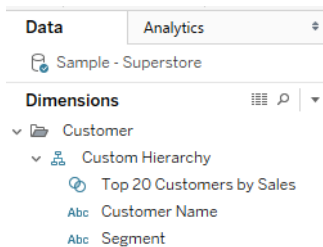
Sets can be used just like dimensions in that if you add a set field to the Columns Shelf or Rows Shelf, the view will be “sliced and diced” by that field. Here is how my scatter plot from earlier looks if I create two separate columns for in and out of the set by dragging the set from the Sets area of the Data pane to the Columns Shelf:



### Within a custom hierarchy

Sets can be used as part of a custom hierarchy in Tableau, which allows you and your end users to easily drill down and back up across different dimensions. For example, you may want to create a hierarchy with customer names that starts with the Top 20 Customers by Sales set, then drills down to individual customer names, then to their segment.

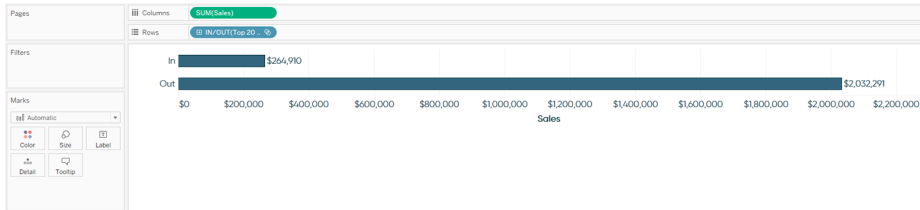
Create a custom hierarchy by selecting the Customer Name and Segment dimensions, right-click, hover over Hierarchy, and choose Create Hierarchy. Once the custom hierarchy is created, you can drag the Top 20 Customers by Sales set into the hierarchy on the Dimensions area of the Data pane and rearrange them to the desired order of the drilldown. At this point, I see this hierarchy in the Dimensions area of the Data pane:



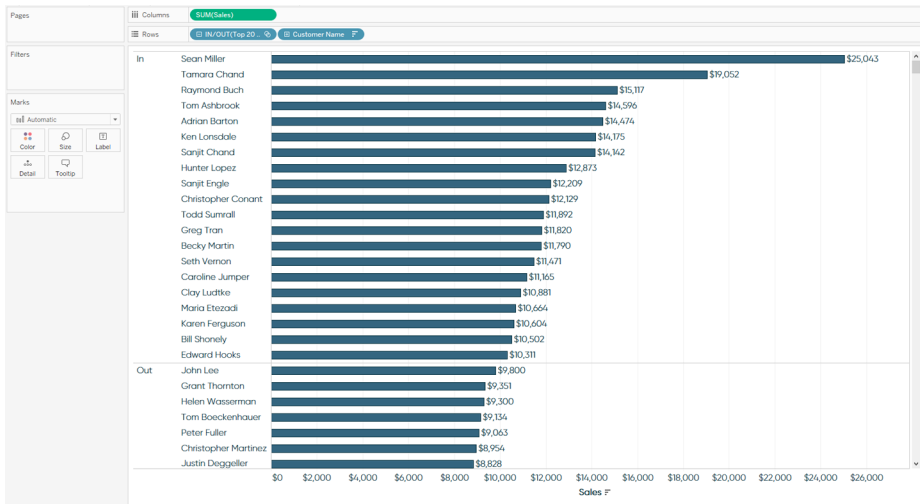
Now that the hierarchy is in place, if I replace the Customer Name dimension with the Top 20 Customers by Sales set dimension in the preceding bar chart example, I will be able to drill down from the top 20 set, to customer name, to

segment. This is accomplished by clicking the “+” symbol on the field(s) on the Rows Shelf.

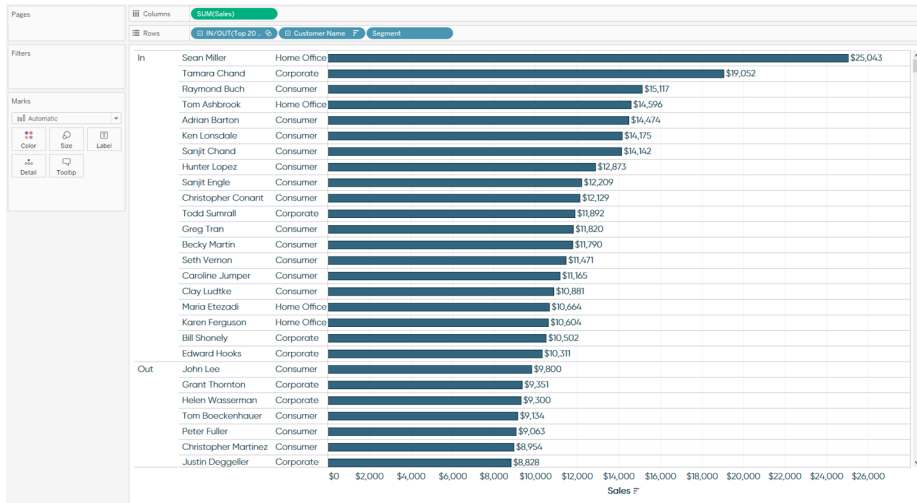
Here’s how the view looks when the bars are aggregated at the highest level of the hierarchy. This view shows sales generated from customers in the set versus sales generated from customers out of the set:



Here’s how the same view looks after clicking the + symbol on the blue pill on the Rows Shelf. We now see sales by whether or not the customer is in our set, but also the names of individual customers:



Lastly, because we built a third level of the hierarchy in for the Segment dimension, we can drill down one more time by clicking the + symbol on the Customer Name pill. This view first breaks sales down by the set that we created, then Customer Name, then Segment:





---

# How to Make a Scatter Plot

After the bar chart and line graph, I find the scatter plot to be one of the most effective visualization options for analyzing data. A scatter plot displays data points at their respective intersections of two measures, and displays every data point on the same view. The marks can then be further encoded by up to three additional measures and/or dimensions by leveraging the Color and/or Size and/or Shape Marks Cards.



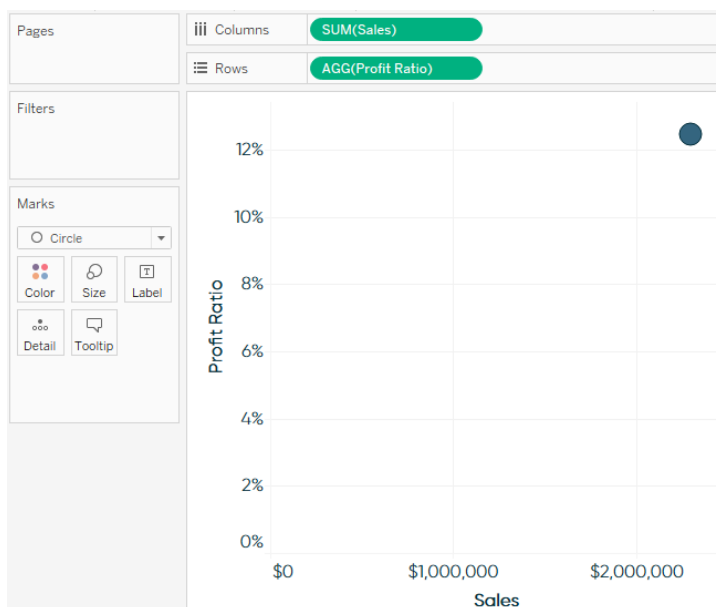
While scatter plots allow you to use several combinations of dimensions and measures, each new encoding increases the cognitive load on your end users. In other words, it makes it harder for them to process the view.

This ability to slice and dice data points in several ways within a condensed space provides an effective means for identifying patterns. Not only that, scatter plots provide a natural way to segment the marks into four quadrants by simply adding a reference line to each of the two axes. You can even take this a step further in Tableau by creating sets for each of the four segments to use for deeper analysis later. This chapter shares how to create a scatter plot in Tableau and use the results to create segments.

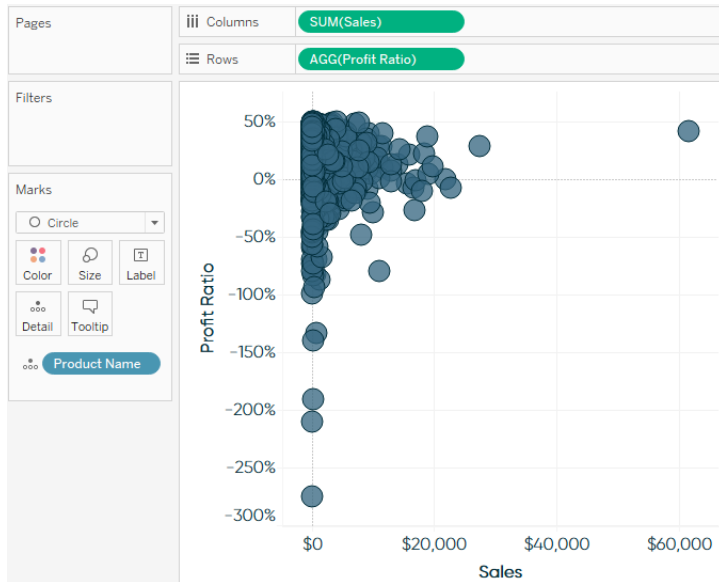
## How to Make a Scatter Plot in Tableau

For this walkthrough, we'll be evaluating all of our products across the Sales and Profit Ratio measures. When you build a scatter plot, one measure will form the y-axis and one measure will form the x-axis. The marks on the view will then be plotted at the intersection of the values on the two axes. It is typically best to put the most dependent metric on the y-axis and the explanatory metric on the x-axis.

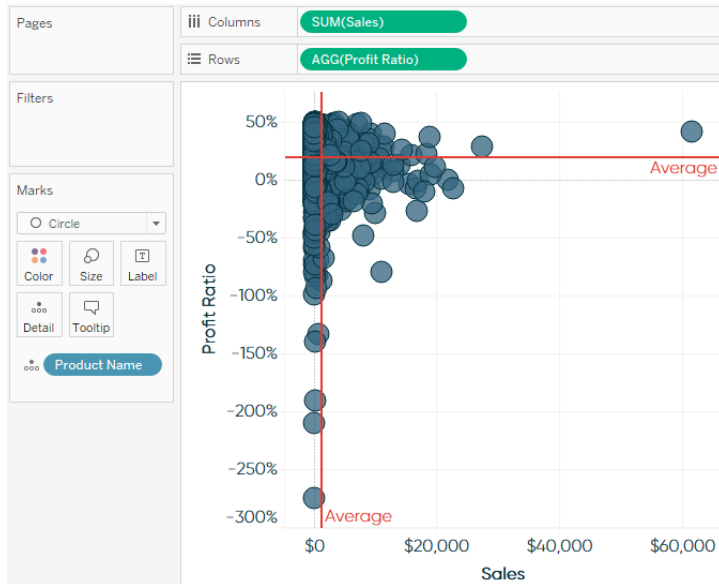
To create a scatter plot, drag and drop the Profit Ratio measure to the Rows Shelf and the Sales measure to the Columns Shelf. Scatter plot is the default chart type in Tableau when two measures are used, so you could have got to this same point by just double-clicking Profit Ratio, then double-clicking Sales to add them to the view. At this point, your view should look similar to this:



So far, we have just one point that represents the intersection of Profit Ratio and Sales for all of the records in our dataset. This is because we have yet to specify a level of detail for our analysis. For more on this topic, you can review [Chapter 10](#). Our analysis is going to look at products, so change the level of detail by dragging the Product Name dimension to the Detail Marks Card. The view has now been changed to this:



What's powerful here is that we are looking at all 1,850 of our products at once, which helps us evaluate them quickly in context of each other. You can see outliers, unprofitable products, and segments are beginning to emerge (i.e., high sales/high profit ratio, high sales/low profit ratio, etc.). One way to make the segments more apparent is to add reference lines to each axis. Here's what the view looks like when I add a reference line for the average of each axis by right-clicking each axis and choosing "Add reference line":



These reference lines create four quadrants on the view that can be used to segment the data:

- **Top-left quadrant:** High Profit Ratio & Low Sales
- **Top-right quadrant:** High Profit Ratio & High Sales
- **Bottom-left quadrant:** Low Profit Ratio & Low Sales
- **Bottom-right quadrant:** Low Profit Ratio & High Sales

This provides some areas to focus on. For example, we don't want to have high sales of products that are causing us to lose money. Also, are there opportunities to increase the sales of our most profitable items? This example created segments using the averages of each measure, but you can easily change the reference lines to the thresholds of your choice for segmentation.

Lastly, these segments or products of interest can be made more permanent by placing them into sets. In this example, our two highest-selling products are also above average in profitability. Let's put these two products into a "positive-outlier" set by selecting them on the view (either through multi-select or dragging a box around them), right-clicking, and choosing "Create set." Once in a set, these positive outliers can be highlighted in different visualizations, which can lead to valuable insights for the business.