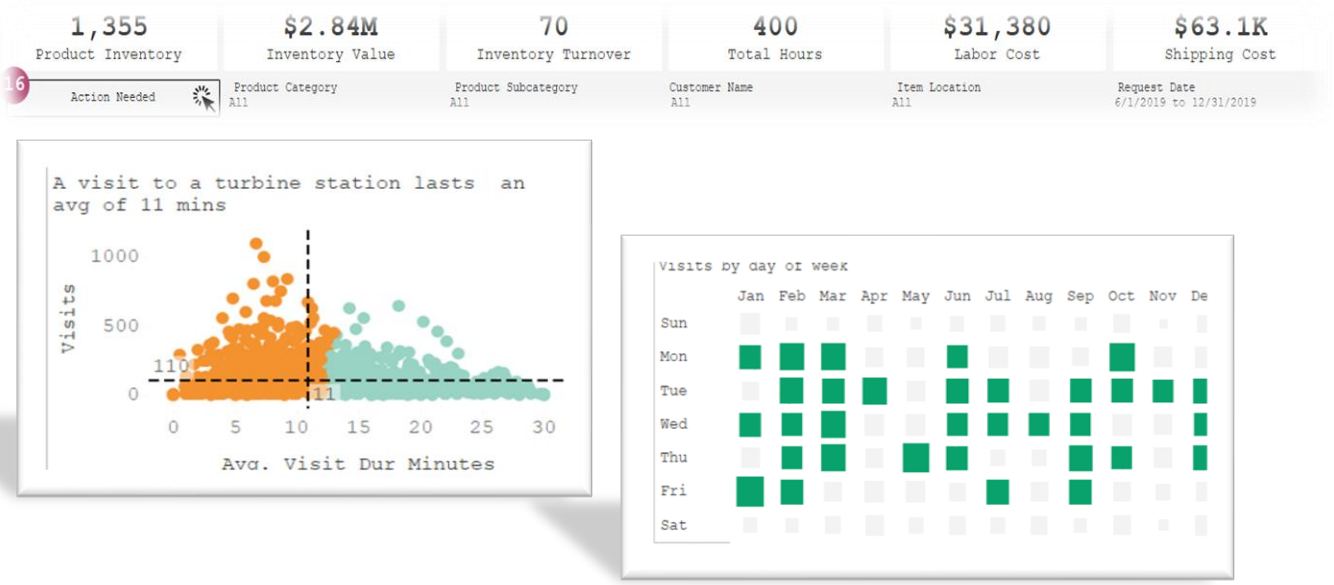
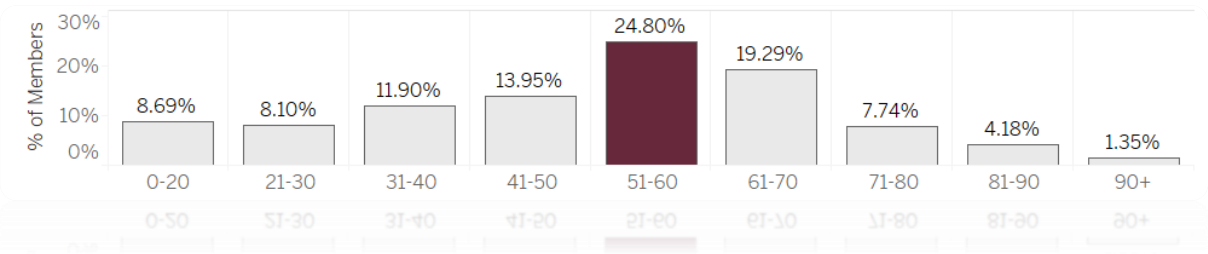


Tableau Like a Pro:

Learning to Use Tableau to Expertly Analyze Data



RYAN NOKES

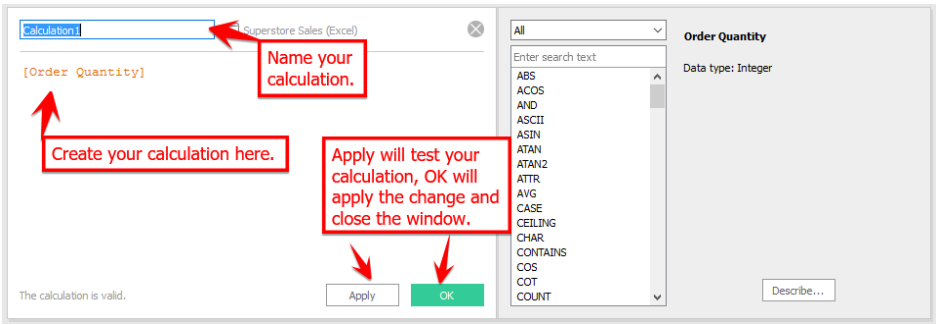
Table of Contents

Calculated Fields

Like any analytical software, whether that's Excel, SAS, R, or Tableau, calculations are the key to going from beginner to advanced. When you master calculations, you master the software tool you are using, and you become an infinitely better analyst. Understanding how these work is crucial to making Tableau work well for you.

In Tableau, when you want to create a calculation:

1. Right click on a particular Dimension or Measure and choose Create > Calculated Field or click the drop-down arrow next to Dimensions.
2. Name your new calculation by writing over the box that says Calculation #.
3. In the window below, write your calculation.
4. Click Apply to test the effect of your calc and click Ok to apply the changes and to close out of the calculation window.



The next several sections will discuss a variety of calculation types as well as specifics on how to use Tableau's built in functions. While it won't go over every function possible, it will cover the main ones. To learn about the others, use the menu along the right-hand side to see explanatory info on how the functions work as well as an example.

Aggregate Calculations & Ratios

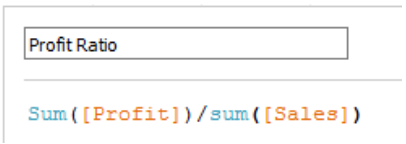
Let's say you wanted to create a ratio of something. Perhaps it's Profit Margin (Profit/Sales) or Revenue per Day (Revenue/# Days). Let's also say that the lowest level of granularity in your data set is at the customer level. If you create a ratio calculation at the lowest level of grain in your data, it will work appropriately. If customer A has Sales of $1000, and Profit of $500, it's a 50% profit. However, when you aggregate at a higher level – say all customers in a particular region – this calculation will break down. Why? Because Tableau is adding the percentages for each customer in the region and then averaging those percentages. So, if customer A has a profit margin of 50%,

customer B has a margin of 60%, and customer C has a margin of -25%, Tableau will show you a profit margin of 85% (50+60-25). That's incorrect. We need it to sum all the sales and all the profit and then take the ratio of that.

Example:

Customer	Sales	Profit	Profit Margin
Customer A	1000	500	50%
Customer B	500	300	60%
Customer C	100	-25	-25%
Total	1600	775	775/1600= 48% - CORRECT Not (50+60+(-25))/3 = 28%

To sum all of the profit figures as well as sum all of the sales figures and then divide by the totals, the calculation looks like: Sum([Profit])/Sum([Sales]). Tableau now knows to sum the figures first and then calculate the ratio, rather than sum all the individual ratios.



To create ratios in Tableau, ensure you use the level of aggregation you need in your calculation. Typically, this will be a sum, but can also be an average, or min/max values.

Logical Calculations

If/Then

An If/Then statement allows you to check whether certain conditions are true. It also gives you the ability to create groups or pull out certain fields from a dimension. They are extremely useful, but using too many with too large of a data set can cause your workbook to slow down.

IF condition1 THEN x ELSE IF condition2 THEN y ELSE z END
You can add as many ELSE/IF conditions as you want. Here's a specific example:
If [Ship Time] > 3:00pm THEN "Late Afternoon or Evening"
ELSEIF [Ship Time] > 12:00pm THEN "Mid Afternoon"
ELSE Morning
END

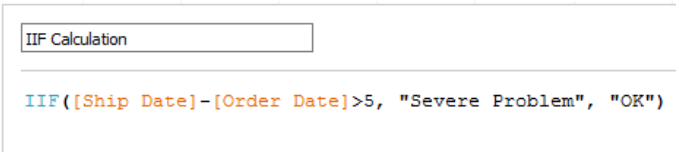
In this case, the first condition will pick up anything that occurs after 3:00pm. The next condition will pick up anything between noon and 3. And the last condition will pick up everything else.

IIF

IIF is very similar in terms of what it does as the IF statement above, but it's syntax is more similar to an IF statement in Excel. You can use either statement.

IIF(test, then, else, [unknown])

IIF([Ship Date] - [Order Date] > 5 (days), "Severe Problem", "OK")



Case/When

Case functions are only applicable to string fields (i.e. text). They check whether certain conditions are met within the dimension or parameter and return a value specified by you.

CASE [Month]
WHEN "January" THEN "Jan"
WHEN "February" THEN "Feb"
WHEN "March" THEN "Mar"
Etc…
END

These can also be used with parameters. The statement below enables you to select which dimension you want to show in a particular chart. The setup of the chart remains the same but the information shown can be changed.

CASE [Choose Dimension Parameter]
WHEN "Marketing Channel" THEN [Marketing Channel]
WHEN "Customer Segment" THEN [Customer Segment]
WHEN "Product Category" THEN [Product Category]
END

IfNull

This function checks whether a field is null, or has no data in it. The first field will be the field you want to use if it isn't null. If the field is null, that's the second condition. But, there's one caveat: the data types have to match between the first and second condition. If the first field is a date, the second field must be a date. If the first field is text, the second must be text, and so on.

IfNull([Order ID],0)

- This checks whether the product has a ship date. If it does, it will return the ship date; if it doesn't have a date, it will return "Hasn't Shipped Yet."

IsNull

This is simply a true/false statement that checks whether a record is null.

IsNull([Zipcode])

- This checks whether the zip code is null or not. If you put Number of Records next to it, you can see how many are null vs. how many are populated. If you use it on the Filter shelf, you can filter out nulls or show only those with nulls.

You can also combine this with the IF statement for a more advanced filter.
If IsNull([Ship Date]) THEN "Hasn't Shipped Yet" End

ZN

The ZN function checks whether measures are null and converts the null values to a zero. So, imagine the following data:

Customer ID	Amount Paid
0123	$500
4567	$2500
8998	
0009	$3000

The blank field is a null value and will be converted to a zero.
ZN's must be aggregated.

```
ZN(sum([Amount Paid])
```

Practice Exercises

Using calculations in Tableau is where you truly start to take your analysis to the next level. Finance and Marketing both want to see the Profit Margin applied to various analyses you've created. They would also like some of the charts to provide more instant insight. We'll use calculations to get those.

1. Create a calculated field with the following formula SUM([Profit])/SUM([Sales]).
2. Name it Profit Margin.
3. Go to the Profit by Marketing Channel worksheet.
    1. If you don't have it, bring Marketing Channel to Rows and Profit to Columns.
4. Add Profit Margin onto the Columns shelf.
5. Turn on labels using the Labels shelf.
6. Right click on the labels on the profit margin side of the chart and choose Format.
7. In the formatting pane at the left, ensure that the Fields drop down menu is set to Agg(Profit Margin).
8. Change both the axis and the chart labels to be percentages, 0 decimal points.
9. Bring out another instance of Profit in between Profit and Profit Margin on the Columns shelf.
10. On this pill you just dragged out, click the drop down arrow and create a Quick Table Calc for Percent of Total. You'll now have three charts side by side.
11. Sort the Profit Margin in descending order. Determine some areas where you could have some quick wins, like Social Media.
12. Rename the tab "Profit Margin by Marketing Channel."
13. Create a new calculated field. This will combine both a date calculation and a case statement. Type in the following calculation:

```
CASE Datename('month',[Order Date])
WHEN 'January' THEN 'Winter'
WHEN 'February' THEN 'Winter'
WHEN 'March' THEN 'Spring'
WHEN 'April' THEN 'Spring'
WHEN 'May' THEN 'Spring'
WHEN 'June' THEN 'Summer'
WHEN 'July' THEN 'Summer'
```

```
WHEN 'August' THEN 'Summer'
WHEN 'September' THEN 'Fall'
WHEN 'October' THEN 'Fall'
WHEN 'November' THEN 'Fall'
WHEN 'December' THEN 'Winter'
END
```

1. Name the calc "Seasons." This is saying that if the text version of the month equals the following names then classify them according to their appropriate season.
2. Now find the Order Qty by Day chart we created. It probably hasn't yet been used in a dashboard so you'll need to go to the worksheet directly.
    a. If you don't have it, bring Order Quantity to Rows.
    b. Bring Order Date to Columns.
    c. Click the drop down arrow on Order Date and set it to be a continuous Month (May 2015) from the second section of dates in the menu.
3. Now, change the Order Date pill from a continuous (green) day to a discrete (blue) Week Number, accessed via the sideways arrow in the first date section of the drop down, where it says "More."
4. Drag your Seasons calc onto Color.
    a. If you can't find where your Seasons calc went, use the magnifying glasss up next to Dimensions.

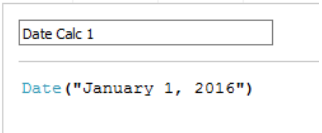
Date Calculations

Date calculations will likely be one of the most common calculations you use. Following is a list of different types of date calcs and when and how to use them.

Date

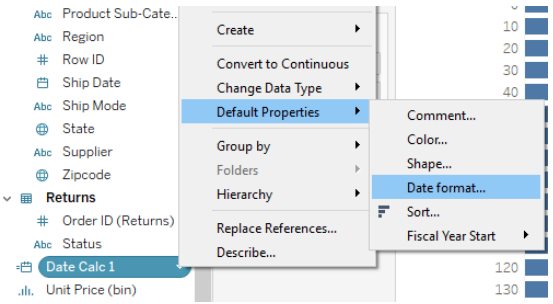
Converts an input – usually a string value – into an actual date that you can then use as dates in your views or in date calculations.

- Date("January 1, 2016")



    - This will convert your date to something like this: 1/1/2016

- Once you click OK you will see the calculated field on the left side of your screen under Dimensions or Measures as appropriate. You can always change it if you need it to be the other.
- You can format this to show however you want by clicking the down arrow on the calculated field and choosing Default Properties > Date Format



The date in this calculation we just created doesn't have to be a hardcoded value (i.e. Jan 1, 2016). This can be another field that looks like a date but is actually a string. E.g. [Year] (type: string or text value). The example below shows how to convert strings or text into dates.

- Date(str([Year]) + '/' + str([Month]) + '/' + str([Day]))
  - This converts three separate text fields into one actual date, E.g. 1/1/2016

DateAdd

Adds a specified time period to a date. You specify the date part you want to add and this calc adds it to the date you've selected. This is useful when you need to calculate new dates, set time thresholds, create reference lines, or create new dimensions.

- DateAdd(Datepart, Increment, Date) => DateAdd('month',3,[Order Date])
  - In the example above, we've added 3 months to the original Order Date.
  - The datepart portion is always in lower case and in single quotes. The table below shows allowable values)

Datepart	Values
'year'	Four-digit year
'quarter'	1 - 4
'month'	1 – 12 or "January", "February", etc.
'dayofyear'	Day of the year; Jan 1 is 1, Feb 1 is 32, Dec 31 is 365
'day'	1 - 31

‘weekday’	1 -7 or “Sunday”, “Monday”, etc.
‘week’	1 - 52
‘hour’	0 -23
‘minute’	0 - 59
‘second’	0 -60

- DateAdd(‘day’, -30, [Today])
    - This subtracts 30 days from the current date of Today. This is useful when you want to show a rolling window of time (say past 30 days).

DateDiff

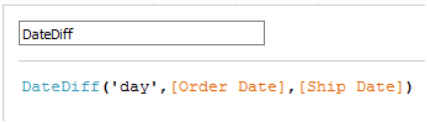
Calculates the time difference between two dates at the level that you specify, such as month, day, hour, etc.

[End Date] – [Start Date] will give you the exact amount of time between the two dates, which sometimes you want. For example, [1/2/2016 8:30 pm] – [1/1/2016 2:45 pm] = -24.79. This is the exact difference in hours and minutes.

Sometimes, you want it at a specified granularity and don’t need as much exactness. Hence the DateDiff function.

- DateDiff(Datepart, Start Date, End Date)
    - Datepart specifies the date granularity you want. See the preceding table for acceptable inputs.
    - Note that the start date occurs before the end date in this calc, so be sure that whatever date you are subtracting from is placed last.

- DateDiff(‘day’, [Order Date], [Ship Date])



DateParse

Converts a string to a specified date format.

- DateParse(format, string)

- DateParse(“dd.MMMM.yyyy”, “January 1, 2016”)

See table below for acceptable formats:

Date string	DATEPARSE function	Display value returned
16.8.97	DATEPARSE (“d.M.yy”, “16.8.97”)	8/16/1997 12:00:00 AM
7/12/91	DATEPARSE (“d/M/yy”, “7/12/91”)	12/7/1991 12:00:00 AM
12Sep2012:9:8:7.6546	DATEPARSE (“ddMMMyyyy:H:m:s.SSSSSS”, “12Sep2012:9:8:8.6546”)	9/12/2012 9:08:09 AM
5:09	DATEPARSE(“h:m”, “5:09”)	5:09:00 AM
97.11	DATEPARSE(“yy.MM”, “97.11”)	11/1/1997 12:00:00 AM
9743	DATEPARSE(“yyMd”, “9743”)	4/3/1997 12:00:00 AM
97-8-14 3:2:1	DATEPARSE(“yy-MM-dd h:m:s”, “97-8-14 3:2:1”)	8/14/1997 3:02:01 AM
19977345	DATEPARSE (“yyyyMdHm”, 19977345”)	7/3/1997 4:05:00 AM

Continued next page

Today()

Returns the current date. This is useful for creating reference lines showing relationships in your data to today (e.g. your deliverables deadlines in relation to today), as well as calculations based on today’s date (such as rolling windows).

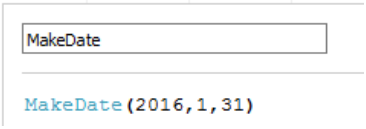
Now()

Returns the current date and time. Same as above, but adds the hours, minutes and seconds. Useful if you need more granular information or are working with intra-day data or data that updates in near real-time throughout the day.

MakeDate & MakeDateTime

Returns a date from a string of numbers, or a string dimension in your data.

MakeDate(2016, 1, 31)
MakeDate(str([Y1][M][D]))
MakeDateTime(#2016,-1-1#, #13:00:00#)



Practice Exercises

To test out our newly acquired calculation skills, let's see if there are differences in avg. time to ship by product line.

1. Add Product Category to the Rows shelf followed by Product Sub-Category.
2. Create a new calculation subtracting Order Date from Ship Date. Use the DateDiff formula and set the time period to 'day'.
    1. Hint: Right click and create a calculated field. If you can't remember the formula syntax, use the scroll box on the right-hand side to find the calc. It will provide an example.
    2. Hint: Datediff calcs work in reverse from a typical subtraction problem, so it's the first date, then the second date.
3. Name it Time to Ship (datediff) and click Ok.
4. Drag out Year(Order Date) to Columns.
5. Drag Time to Ship (datediff) out to Columns. Set it to Average.
6. Also, drag Time to Ship (datediff) onto Color.
7. Set stepped color to be 3 steps.
8. Edit the axis label by right clicking on the axis and renaming the title to be "Avg. Time to Ship in Days."

9. Name the tab “Time to Ship.”
10. Save your work.

Parameters

Parameters are some of the most flexible, and frankly, most awesome features of Tableau. A parameter allows you or your end user to change the content that appears in worksheets and dashboards. This can be something simple like a value or something much more advanced. Here we’ll cover several of the most commonly used examples.

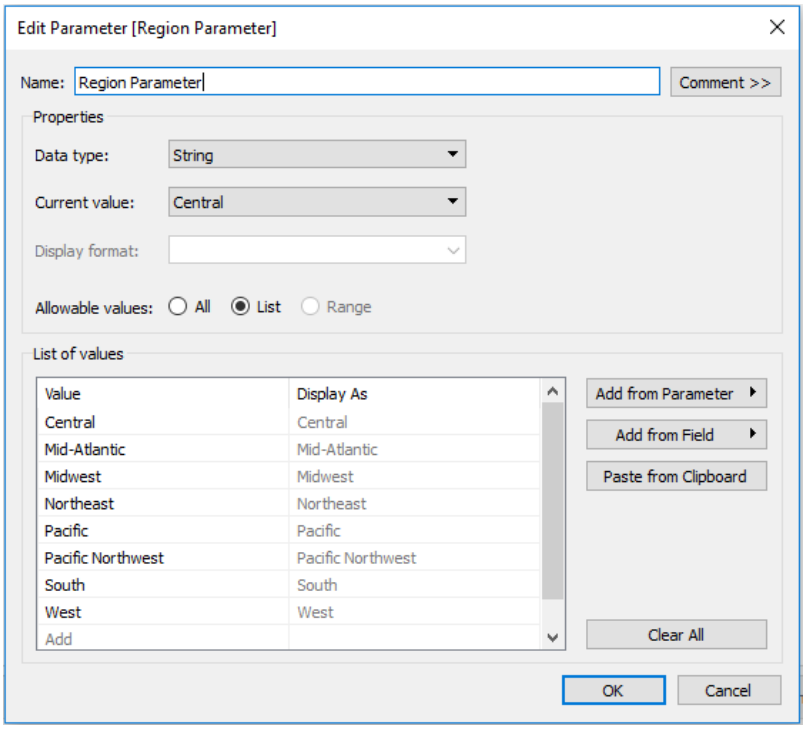
Pre-defined Lists for Faster Filtering

As a general rule, parameters run faster in Tableau than Quick Filters. Every time a Quick Filter runs it pings your data source, which can slow you down. Parameters are retained in memory. So, if you have a list of values you want to use as a filter that you know **isn’t** going to change, or will change infrequently, use a parameter instead.

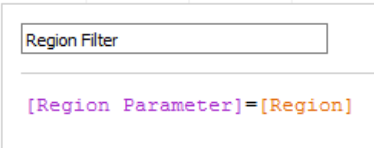
For example, let’s say you have regions: North, South, East, West, MidWest, and Pacific Rim. You know that these are always the regions. You have a Dimension in your data called Region. You could use the Region field as a Quick Filter. However, you could retype these values in a Parameter and use that as a filter instead and gain processing speed. The caveat here is, if your Dimension is going to change frequently, say for example, Customer Name, you don’t want to be reupdating your Parameter all the time. Quick Filters are dynamic, Parameters are static. It’s a limitation of Tableau.

To create the parameter:

1. Right click anywhere in the data pane along the left and choose Create > Parameter.
2. Give it a title.
3. Change the Data Type to “String.”
4. Change the radio button at the bottom for Allowable Values to “List.”
5. Type in the values down below.



6. Choose ok.
7. Right click again and choose Create > Calculated Field. Name your calculation (i.e. Region Filter).
8. Drag the newly created parameter into the calculation window where your cursor is so that it reads [Region Parameter] = [Region] (or whatever names you've chosen).



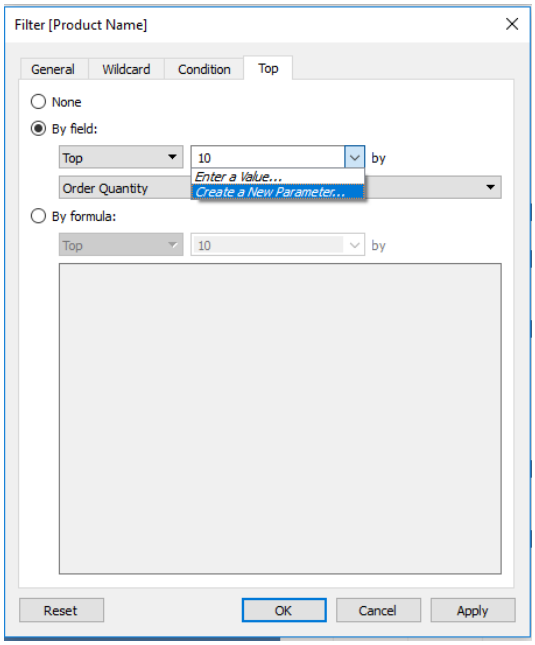
9. Choose ok and then drag your calculation to the Filters shelf. Set the filter to show only True.

Top N Filter

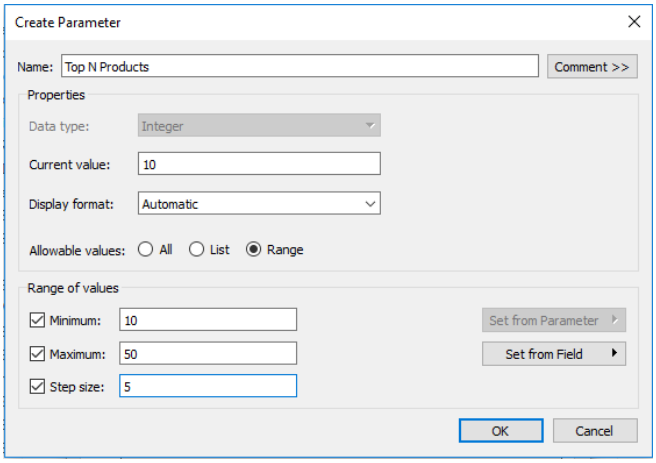
Filtering and sorting are two of the most used tools in any analyst's toolkit. Often you want to see just the top 10 of something. But what if your audience has different needs? Maybe someone wants to see the top 20 or the bottom 20. A parameter allows you to do this.

1. Let's look at top Products by Order Quantity. Drag Product Name to rows and Order Quantity to columns.
2. Now drag Product Name to the filter shelf.

3. A filter box will appear with a number of tabs along the top. Each of those allows you flexibility in how you filter. For now, go to the Top tab.
4. Choose "By Field."
5. In the drop down where it says 10, select "Create a new parameter."



6. Name your parameter "Top N Products."
7. Set your min and max values – in this case let's do 10 and 50. Set the step size to 5 (this is by how much each change increments upwards).

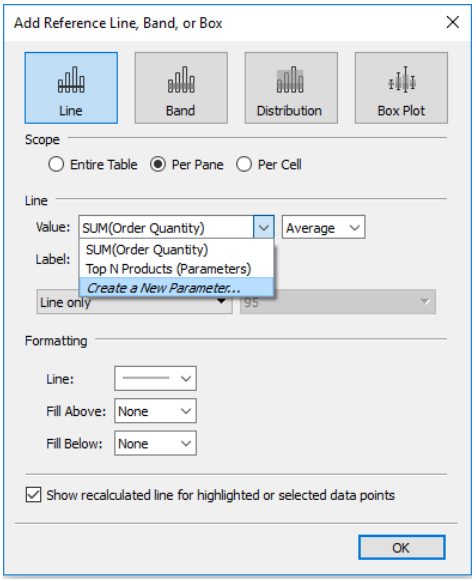


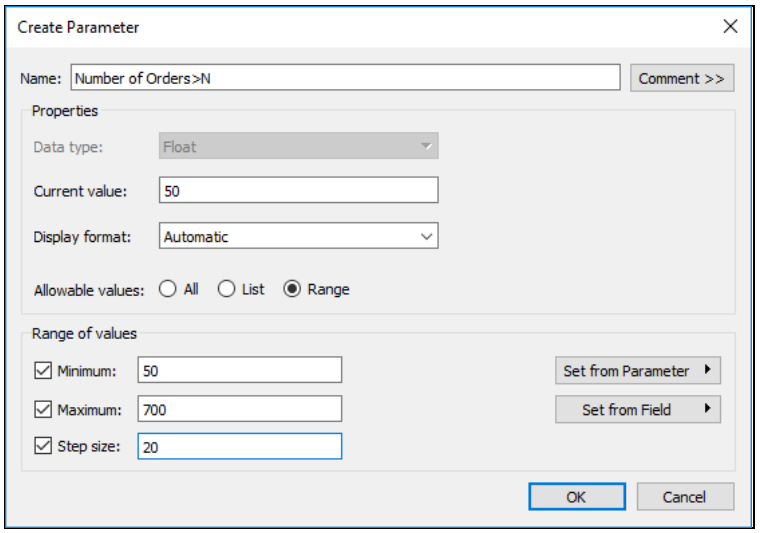
8. Click ok and a small box with a slider will appear on the right. Use the slider to adjust the parameter and explore how it changes the viz.

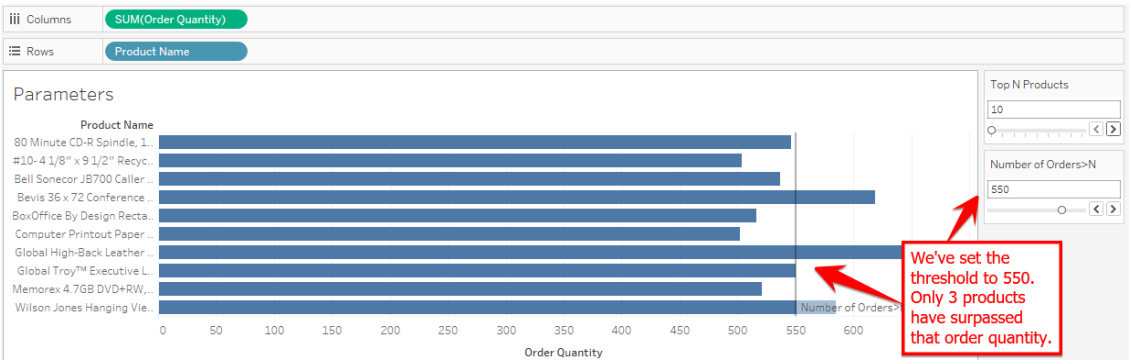
Reference Line Parameter

As mentioned previously, reference lines can be a very beneficial way to highlight key information, show thresholds and targets, highlight outliers, etc. Parameterizing these allows your end user to see the data in relation to the targets or thresholds or reference points they want to see.

1. On the same chart we just created, right click on the axis and choose Add a Reference Line.
2. In the Value section in the middle, click the drop down arrow (where it says Sum(Order Quantity) and change that to "Create a New Parameter."
3. Name this "Number of Orders > N"
4. Make it a range of values in the Allowable Values section.
5. Set the min to 50 and the max to 700.
6. Set the step size to 20.
7. Now play with the slider and see how many fall above the line.



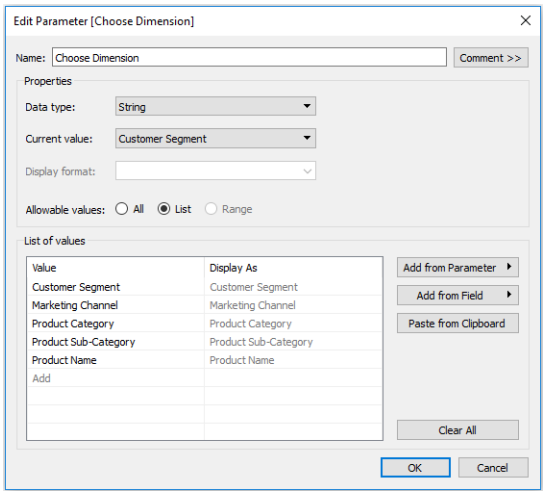




Swapping Dimensions or Measures in a View

We now get into a more advanced use of parameters, but one that is extremely useful. Say you have a well formatted, insightful chart. Rather than recreate the same chart for six different metrics, you'd like your user to be able to see whatever they want to see with just 1 click.

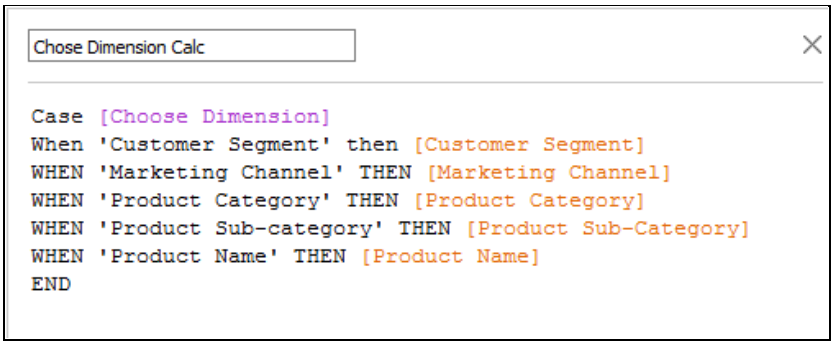
1. Create a new worksheet at the bottom.
2. Right click in the data pane and create a new parameter.
3. Name it "Choose Dimension."
4. Change the Data Type to a String.
5. Set the Allowable Values to a List.
6. In the window down below, type the following field names, one per line.
    - Customer Segment
    - Marketing Channel
    - Product Category
    - Product Sub-Category
    - Product Name
7. Click ok.



8. Now right click and create a calculated field using a Case/When statement. Name it Choose Dimension Calc.

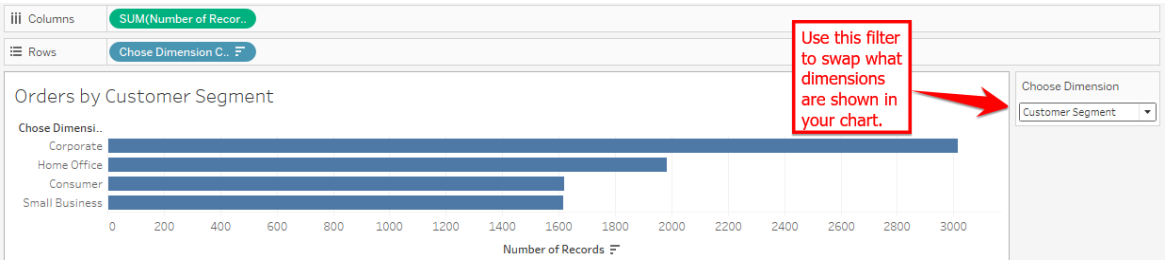
```
Case [Choose Dimension]
When 'Customer Segment' then [Customer Segment]
WHEN 'Marketing Channel' THEN [Marketing Channel]
WHEN 'Product Category' THEN [Product Category]
WHEN 'Product Sub-category' THEN [Product Sub-Category]
WHEN 'Product Name' THEN [Product Name]
END
```

   - Note the following:
     - Capitalization doesn't matter in the function portions in blue.
     - You need the apostrophe marks around the actual text shown in your parameter
     - You can drag in your orange Dimension fields from the data pane.
     - You must close it with an end statement.
     - If you are used to writing SQL, the case statement is obviously something you are familiar with. Note that Case syntax in Tableau is slightly different. It's Case [Field] then your When conditions.



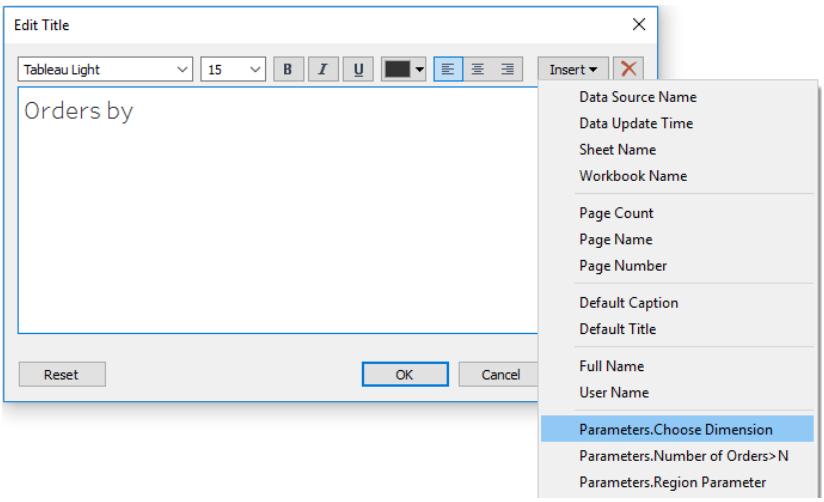
9. Bring this calculation to the Rows shelf and Number of Records to the Columns shelf.
10. Right click on the parameter you created at the bottom of the data pane and choose "Show Parameter Control."
11. Click the Sort Descending icon.
12. Now select different drop downs from the Choose Dimension parameter and see how the chart changes.

You can use this technique to swap out dimensions, swap out measures, swap out headers on your axes in a chart, and change the title. This gives you almost infinite flexibility in how you present and share information.



To illustrate just one more example of this:

1. Double click the title and name it "Orders by ." (Be sure to include the space after "by").
2. Then click the Insert drop down menu at the right.
3. Select your [Parameters.Choose Dimension] field. Click ok
4. Now change the parameter values and see how it changes your title.

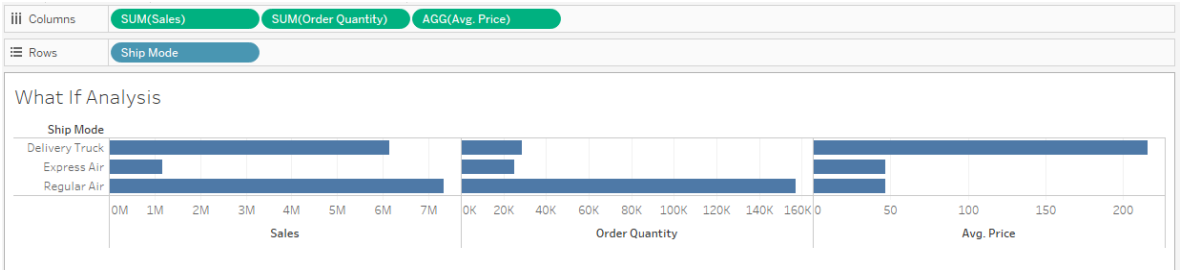


What If Analysis

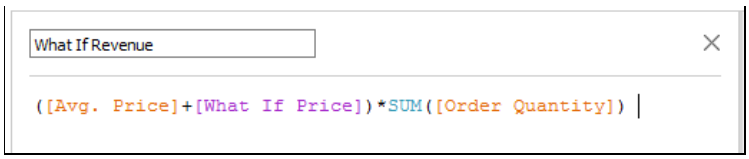
Allowing users to perform a "what if" analysis can be an extremely effective way to motivate action. For example, if you highlight your company's current state, and give them the ability to change some assumptions and values and see where you could end up, people want to make that a reality.

Let's say we wanted to see the average price per Ship Mode.

1. On a new tab, drag Sales and Order Quantity to the columns shelf. They'll both default to Sum, which is what we want.
2. Drag Ship Mode to rows.
3. Create a calculated field showing sum([Sales])/sum([Order Quantity]).
4. Name this Avg. Price and drag it to the right of the other two pills on the columns shelf.
5. Turn on labels for each.

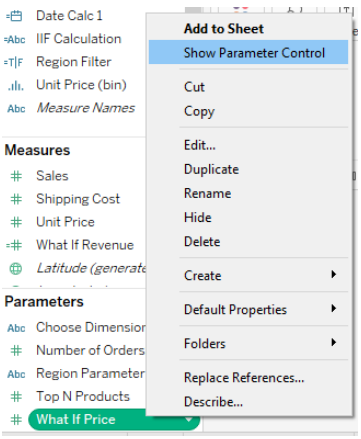


6. Now click on the pill for Avg. Price and then click the Sort Desc icon.
7. Now create a new parameter entitled What If Price.
8. Set the Data Type to Float. Set the Min value to 0 and the Step Size to 10. Uncheck the Max value (this will allow your end user to set the number to whatever they want).
9. Next, create a new calculated field and name this What If Revenue.
10. Type in the following formula: ([Avg. Price]+[What If Price])*SUM([Order Quantity]) (There are duplicate order IDs in the data).
    - This formula takes the average price and allows your end user to add an amount to the average price. Then you multiply the new price by the number of instances of that to determine the overall impact on revenue.



11. Add the What If Revenue pill to the right of the other pills on the columns shelf.

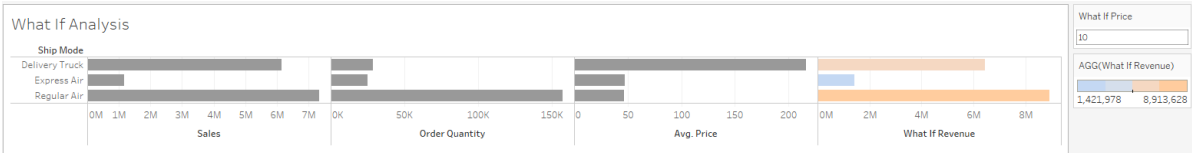
- If the "What If" parameter box does not appear, go down to the Parameters section at the bottom left, find the parameter, right click it and choose "Show Parameter Control."



12. Finally, create a new calculation entitled Forecasted Revenue Increase using the following formula: [What If Revenue]-SUM([Sales]). Drag that to the columns shelf as well.
13. Drag the same calculation to the color shelf on the segment under the marks card that is labeled AGG(What If…).
    - This will convert all the other bars to a gray color and provide a default blue sequential color scheme on the last bars in the chart.

14. Click the down arrow on the legend icon at the top right. Choose Edit Colors.
15. Change the color to Orange-Blue Light Diverging. Set it to Stepped Color with 4 steps, and check the Reversed box.

16. Now, look at the change in the forecasted revenue.



Practice Exercises

Finance wants to be able to have more control of how they look at the numbers. We'll use parameters to give them this flexibility.

1. Go to the tab labeled "Best Customers."
2. If you don't have it, bring Customer Name to Rows and Profit to Columns. Color by Profit. Sort Desc.
3. Right click on the Measures pane and Create a Parameter.
4. Title your Parameter "Top N Most Profitable Customers."
5. Set the Data Type to Integer and set the Allowable Values to Range.
6. Set the range from 1 to 50, with a step size of 2. Click ok.
7. Now right click on the parameter at the bottom left and choose "Show Parameter Control."
8. Create a new calculated field entitled "Index." In the formula window, type Index().
    - We didn't cover this much above but it works very similarly to the Rank calculations. Since we've already got customers sorted by profit in descending order, bringing out the Index calc will show the rank of each customer with 1 being the most profitable.
9. Bring this out to the Rows shelf. It will automatically start as continuous and come after Customer Name.
10. Click on the down arrow and set it to Discrete. Now move it before Customer Name.
11. Create a new calculated field entitled "Top Customers." Type the following formula:

    [Index]<=[Top N Most Profitable Customers]

    - This will find the rank and show everyone before that and up to it.
    - You can find the fields in your date pane and just drag them into the calculation editor.
12. Bring this new Top Customers calculation to the Filter shelf. Set it to True.

13. If you don't want to see the actual customer rank, click on the down arrow on that pill and uncheck "Show Header."
14. Move the slider around to see how the view changes.

Using Actions to Create Interactive Dashboards

Actions provide interactivity that take your dashboard to the next level. They are a great way to conserve space, reduce the number of quick filters, increase workbook speed, and most of all, allow your users to get on-demand details.

There are 3 types of Actions: Filter Actions, Highlight Actions, and URL Actions.

- Filter Actions – allow you to click an item in one chart and filter other charts (providing drill down detail in one view)
- Highlight Actions – allow you to select a value and see that value's position in other charts (allowing you to focus on what matters and avoid noise)
- URL Actions – allow you to insert a url link that takes you to another destination (could be a webpage or could be another dashboard on Tableau Server)

Let's look at examples of all 3.

Filter Actions

1. First, let's build a very simple dashboard. Open a new dashboard tab by clicking the tab at the bottom or right clicking.
2. Create a heatmap in a worksheet with Customer Segment on rows, Regions on columns and colored by Sales (see below). Bring this out onto the dashboard canvas.
3. Drag table with subtotals that we created out underneath it. Eliminate the legends and filters that appear on the right hand side.
4. Ensure the heatmap is set to Entire View and the text table is set to Fit Width (from the drop down menu at the top underneath the Help menu).

5. Double click the title for the top chart and rename it to “Sales by Customer Segment per Region.” Double click the title for the bottom chart and rename that to “Sales and Profit Detail.”

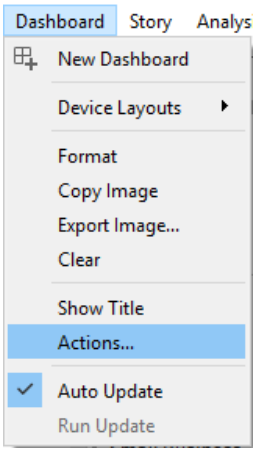
Sales by Customer Segment per Region (click customer segment to see more details)

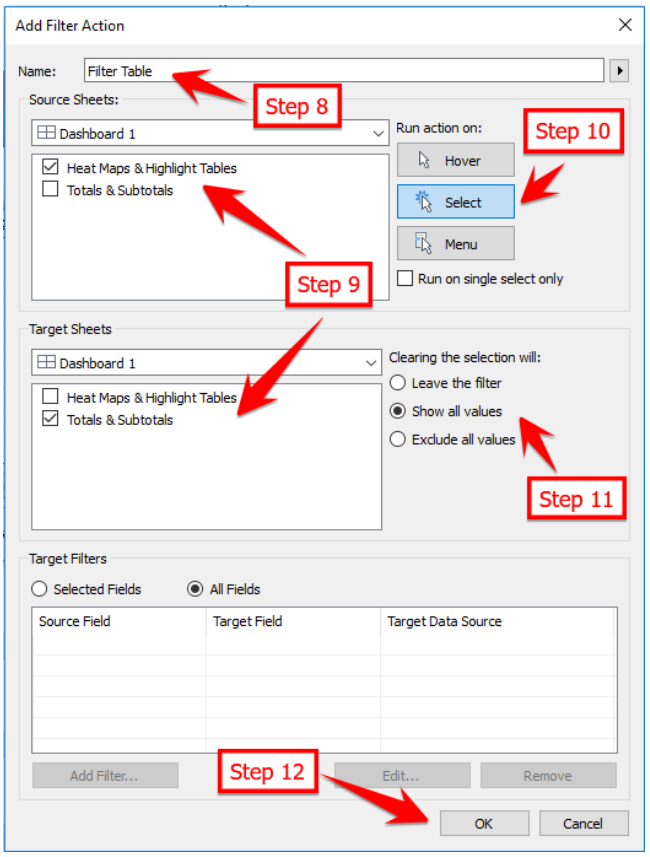
Customer Segment	Region							
	Central	Mid-Atlantic	Midwest	Northeast	Pacific	Pacific Northwest	South	West
Consumer	249,257	349,914	562,060	482,892	68,848	154,883	675,820	474,700
Corporate	468,250	487,355	1,203,828	729,364	119,649	220,224	1,226,136	955,110
Home Office	220,919	277,298	667,069	596,968	90,635	131,304	821,953	661,145
Small Business	215,424	312,269	548,846	394,091	40,059	109,900	656,261	486,631

Sales and Profit Detail

Customer S..	Marketing ..	Product Catego..	Profit	Sales
Home Office	Direct Mail	Furniture	11,608	94,055
		Office Supplies	765	33,625
		Technology	1,557	53,048
		Total	13,930	180,729
	Direct Sales	Furniture	20,502	334,583
		Office Supplies	3,768	92,307
		Technology	55,260	284,520
		Total	79,531	711,409
	Email Marketing	Furniture	7,801	38,547
		Office Supplies	129,582	640,458
		Technology	75,732	526,200
		Total	213,114	1,205,204
	Google Adwords	Furniture	22,977	142,300
		Office Supplies	-21,980	106,773
		Technology	13,662	279,826
		Total	14,659	528,899
	SEO	Furniture	125	13,435
		Office Supplies	4,774	37,897
		Technology	3,097	14,241
		Total	7,996	65,573
	Social Media	Furniture	9,176	30,165

6. Click the Dashboard file menu and then click Actions.
7. Click Add Action in the bottom left corner. Choose "Filter."
8. A new box will appear. Rename this to something descriptive.
9. Uncheck the Totals & Subtotals tab box in the top pane, and uncheck the Heatmap tab box in the bottom pane.
    - This top pane enables you to say which charts will filter other charts. Leaving them both checked will mean that clicking in either chart will filter the other. We want to only filter from the top chart down to the bottom one, in this case.





10. Set the "Run Action On" option to Select.
    - This tells Tableau when to make your action work.

- "Hover" filters the other charts simply by mousing over a portion of the chart.
- "Select" filters it upon a click.
- "Menu" will pop up with a blue hyperlink in a tooltip when you mouse over prompting you to click it if you want to see more information. The text of the hyperlink is determined by what you title the filter action.

11. Change the "Clearing the Selection Will" to Show All Values.
    - "Leave the Filter" will leave the charts filtered to whatever your last selection was and will remain so until you click to filter on something else.
    - "Show all Values" will revert back to showing everything as soon as you deselect your selection by clicking again.
    - "Exclude all Values" will blank your chart out until something else is clicked. This is an extremely useful feature for data on demand and will be discussed in an example later.
12. Click ok.
13. Now click different squares in the heatmap and notice how it filters the table down below.

Highlight Actions

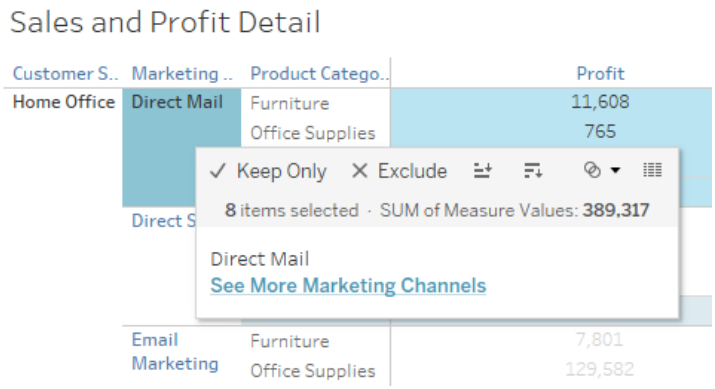
Follow the same procedure for Highlight Actions, except choose Highlight instead of Filter. Note that you can use multiple actions on the same charts. So you can use both a filter action and a highlight action on the same charts.

URL Actions

To demonstrate a URL action, let's build the following dashboard:

1. Create a new dashboard.
2. Bring out the Totals & Subtotals text table that we named Sales and Profit Detail.
3. Add an action, but this time set it to URL when you click the Add Action button.
4. Name it "See More Marketing Channels."
5. Check the box for Totals & Subtotals.
6. Set "Run Action On" to Menu.
7. Paste the following link into the URL box: http://www.digitalinformationworld.com/2015/09/infographic-the-19-paths-to-extraordinary-growth.html
8. Click ok, and then click on a bar in the top chart.

9. Click the link.



Since pretty much anything and everything is on the web or can be put on the web, this is a very flexible and powerful option. While we won't cover this here, you can also include parameters in these links to pass filters (i.e. to filter one dashboard to another that's not in the same workbook) or to see certain values on a webpage.

Practice Exercises

The Logistics group would like to be able to drill down to more details from your initial analysis. And the Marketing group would like to be able to compare our results vs. industry best practices. We'll add this insight with actions.

1. Go to the Time to Ship chart we've been working on. Add Ship Mode to the Row shelf in front of Avg([Time to Ship]).
   - If you don't have it, put Order Date on Columns, followed by Avg(Time to Ship]).
   - Place Product Category and Product Sub-Category on Rows.
   - Place Avg(Time to Ship]) on Color.
   - Set Stepped Color to 3 steps.
   - Rename the tab "Shipping Time Analysis."
2. Create a new dashboard entitled Shipping Times.
3. Bring out just the Time to Ship chart.
4. Move the legend to the bottom by dragging it underneath the chart. Resize to make it fit nicely.
5. Click on the chart, then click on the Analysis menu, and add a Quick Filter for year.
   - It might get placed below your legend at the very bottom. Click and make it a floating element and move it up to the top.

- Set it to be a Multi Value Dropdown type.

6. Now, go to the Shipping Analysis Dashboard. We're going to add a few actions.
7. Go to the Dashboard menu. Click Actions > Add Action > Filter.
8. Check only the first box in the top pane. Uncheck the first box in the bottom pane.
9. Set "Run Action On" to Select. And set "Clearing the Selection Will" to "Show all values." Click ok.
10. Now Add Action (again) > Filter. Title this one "Drill down to Shipping Time Analysis."
11. Leave the first two boxes in the top pane checked.
12. Set Target Sheets in the bottom pane to Shipping Times (the dashboard we just created above).
13. Set Select on the "Run Action On" to Menu. And set "Clearing the Selection Will" to "Show all values." Click ok.
14. Now click the bubble chart and see how the values change. Then click the blue link in the tooltip. Notice how it filters that chart.

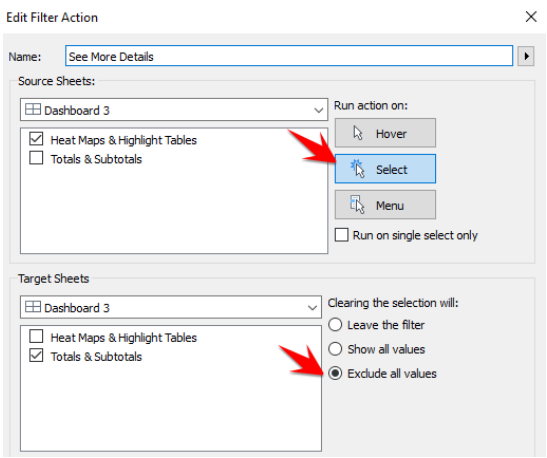
Combining Parameters and Actions: On Demand Charts

Often times you want to have a chart that provides more detail, but you don't want to overwhelm a dashboard with too much information all at once. This is where having a chart or table pop up on demand can be extremely beneficial.

In order to make this work, we must make use of layout containers. These will be covered in the Advanced Formatting section. For now, it's enough to know we need layout containers to make this work.

1. Create a new blank dashboard.
2. From the left hand pane, find the label that says "Vertical" with the two hollow stripes icon. Drag that out until you see a gray shaded box that runs the entire length of the dashboard.
3. Bring out the heatmap that we entitled "Sales by Customer Segment and Region" and set to entire view.
4. If you can't see the bottom of the chart, scroll all the way down, so that you can. Now, drag the text table entitled "Totals and Subtotals "Sales and Profit Detail" into that vertical layout container. You'll know you have it right when you see a dark blue outline and a gray shaded space inside.

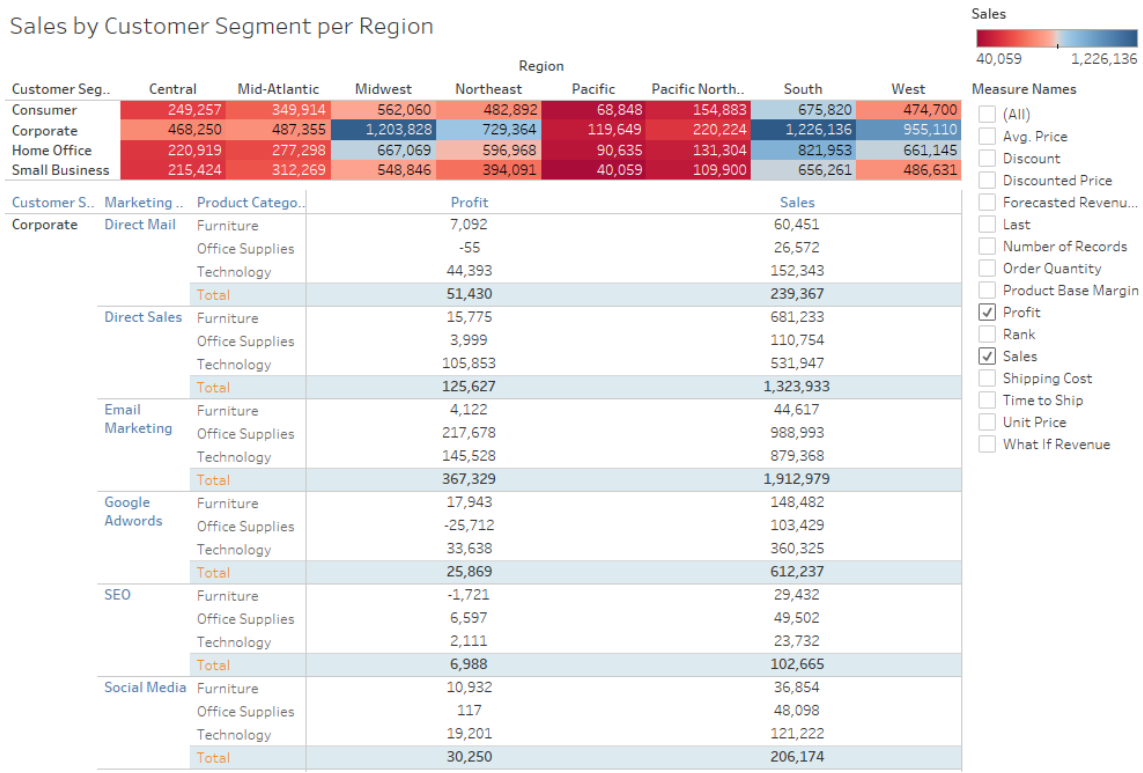
5. Now click on the bottom chart and then set this to Fit Width in the drop down box on the tool bar.
6. Right click on the title of that table and select "Hide Title."
7. Now, go back to Dashboard > Actions and add a Filter Action.
8. Name this "See more details…"
9. Then in the first pane, check the box for the heatmap and uncheck the Totals and Subtotals box.
10. Set "Run Action On" to Select.
11. Under "Target Sheets" down below, check only the Totals and Subtotals Formatted box.
12. Set the radio button to "Exclude all Values."



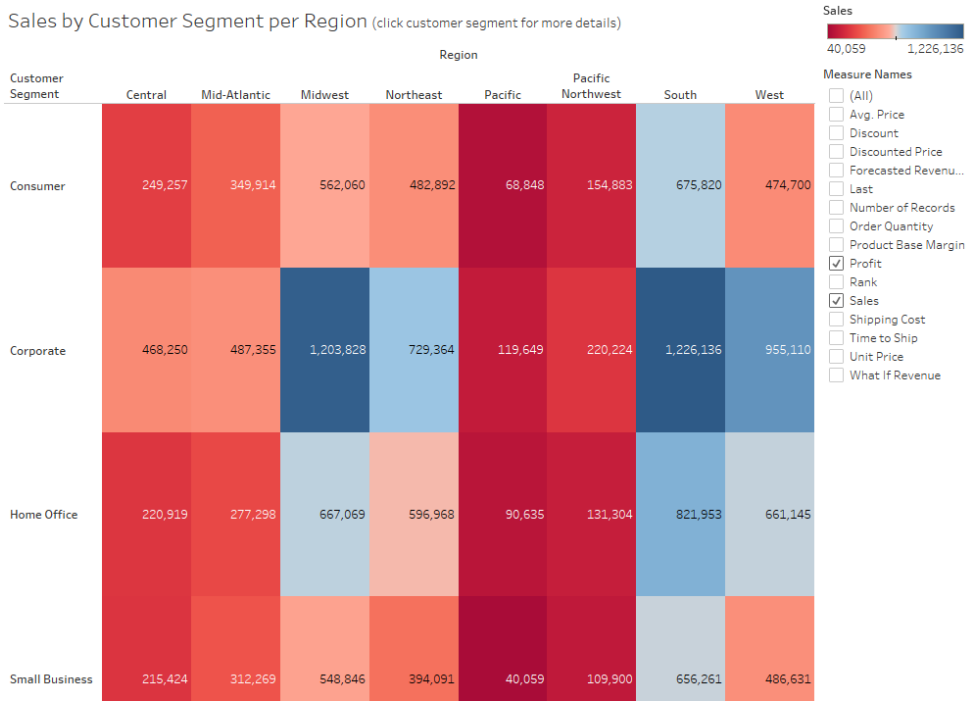
- This last step is crucial to have the chart disappear, and only reappear when desired.

- When you click a chart and see the tooltip pop up to see more details, you can click that to have the table down below appear. When you click somewhere else, that table will disappear again.

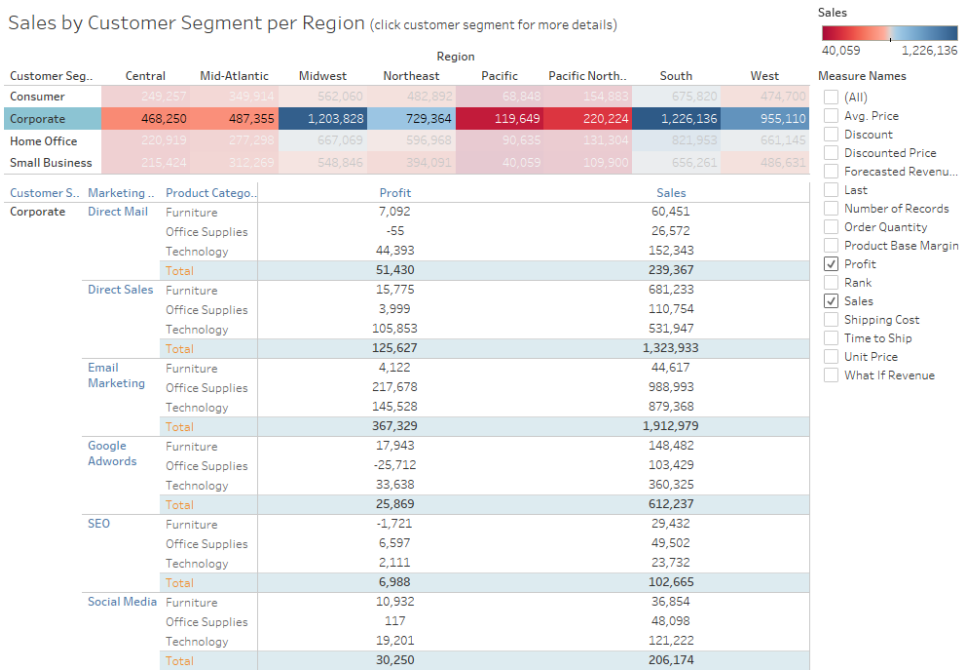
13. Click ok. Your view will look similar to this:



14. Click on one of the Customer Segments in the top chart. Now, click it again. You'll notice that the bottom chart disappears. This is what we want.
15. Add in the following parenthesis after the title of the top chart: (click customer segment for more details). Set the font size on that to 10.
16. If just the bottom chart disappears, leaving a lot of white space, click into that area and you'll see a push pin icon at the top left of the chart. Click that to uncheck it. The push pin fixes a particular size, which then breaks the pop up effect. We want to ensure the charts are set to automatic, not fixed sizes.
17. Now click back into the top chart. Your viz should look like this:



18. Click the chart once more to see how the details pop up. Feel free to change the Measure Names filter on the right to get a sense of how it affects the chart down below.



Advanced Formatting & Dashboard Best Practices

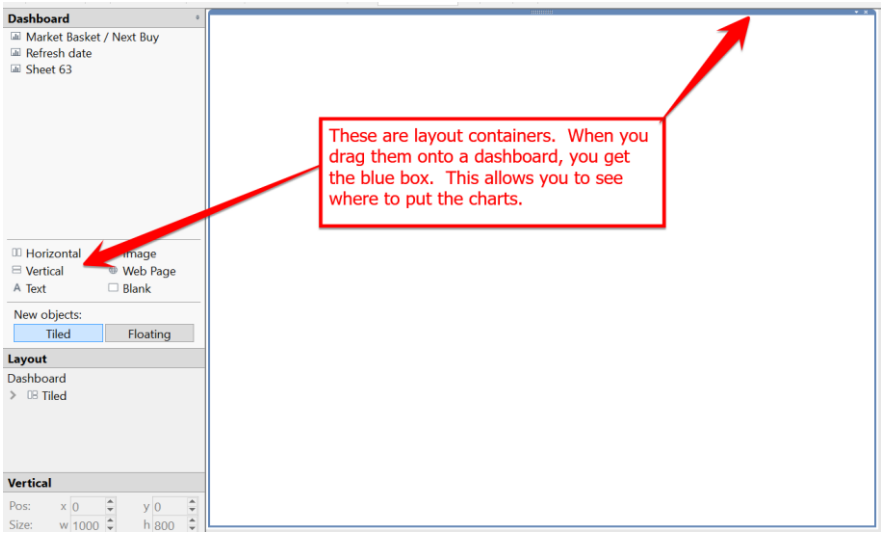
Layout Containers

Layout Containers control the spacing between dashboard components. They allow you to format common elements and move multiple dashboard objects at the same time.

Layout Containers come in one of two options: Horizontal or Vertical.

- Horizontal - groups worksheets and dashboard components left to right across your page and allows you to edit the height of all elements at once.
- Vertical - groups worksheets and dashboard components top to bottom down your page and allows you to edit the width of all elements at once.

Layout containers correspond with the button in the middle of the left side Dashboard panel labeled "Tiled." This is the Tableau default option.



Floating Elements

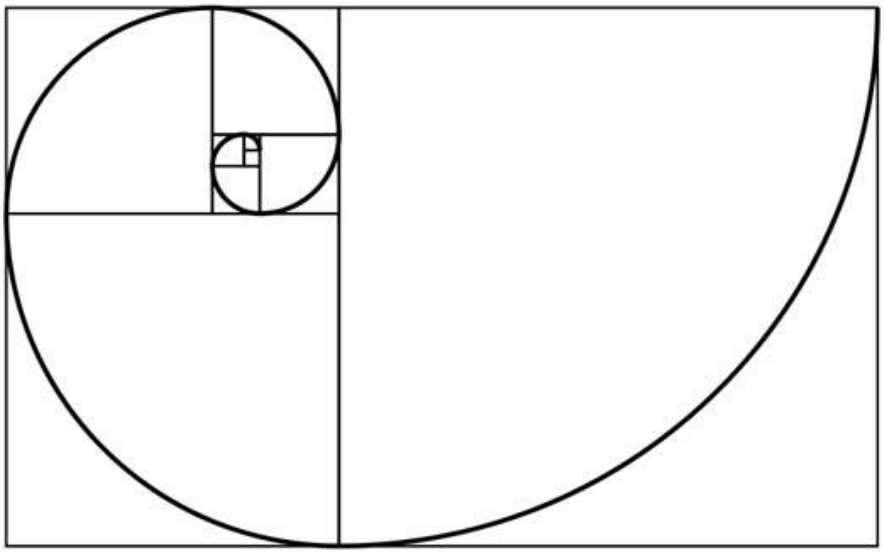
Floating elements are exactly that: elements that float exactly where you want them on your dashboard. They can overlay other charts or dashboard elements, such as a color legend next to a map (in the blank ocean part). Floating allows you to control the exact spacing between objects as well as the heights.

When to Use Which

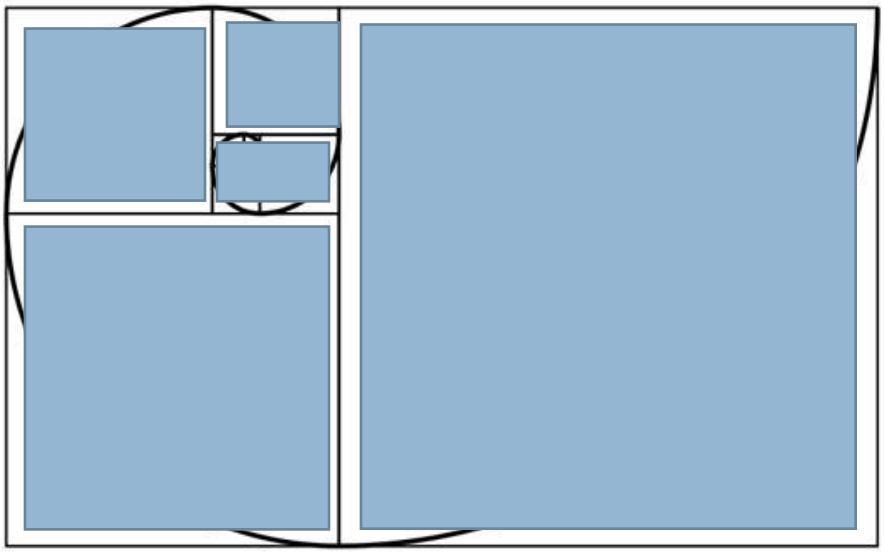
Layout containers are good for sizing and spacing multiple items at once, saving time. Floating is good to emphasize certain areas or layer concepts in one dashboard. Floating is also good if you want exact control of your formatting.

Effective Dashboard Layouts

The Golden Ratio (or Fibonacci Sequence), is a ratio found in patterns we see all around us, and have been found to be aesthetically pleasing. Here's a sample of harmonic spacing using the Golden Ratio:



We can use the same layout to help us prioritize the content and message we want to share in our dashboards. Here's an example of some chart positioning overlaid.



Layout Best Practices

1. Emphasize the most important thing by making it largest or stand out in some other way.
2. Your eye starts at the top left, so put a key insight there, or a key navigational element that will guide them through the rest of the dashboard and lead them to the conclusion they need to find.
3. Use floating elements to focus formatting and cluster elements (like legends) with their associated charts.
4. Use tiled layout containers to create uniformity, and size and move things together at the same time.

Titles and Labeling

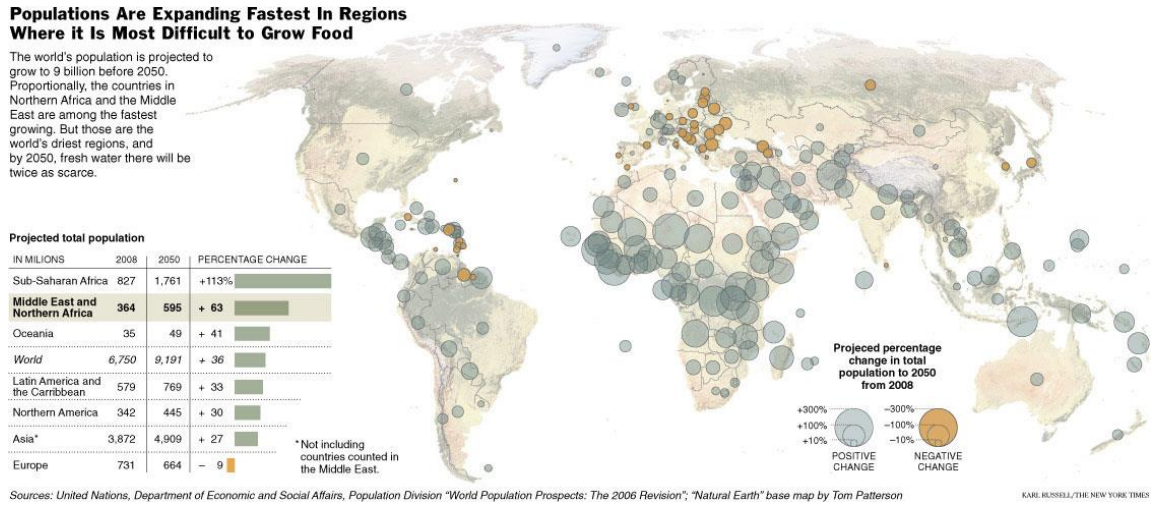
Look at enough infographics in the New York Times and you'll start to see some commonalities:

- They always include descriptive titles, often with the key insight right there.
- They often use subtitles to provide further clarity or details, or highlight a second key insight.
- They use call outs or tooltips to highlight key details in the actual view.
- They provide instructions on how to interpret the chart, as well as how to use the chart (i.e. click here to see more detail).
- They provide descriptive axes on charts and legends.

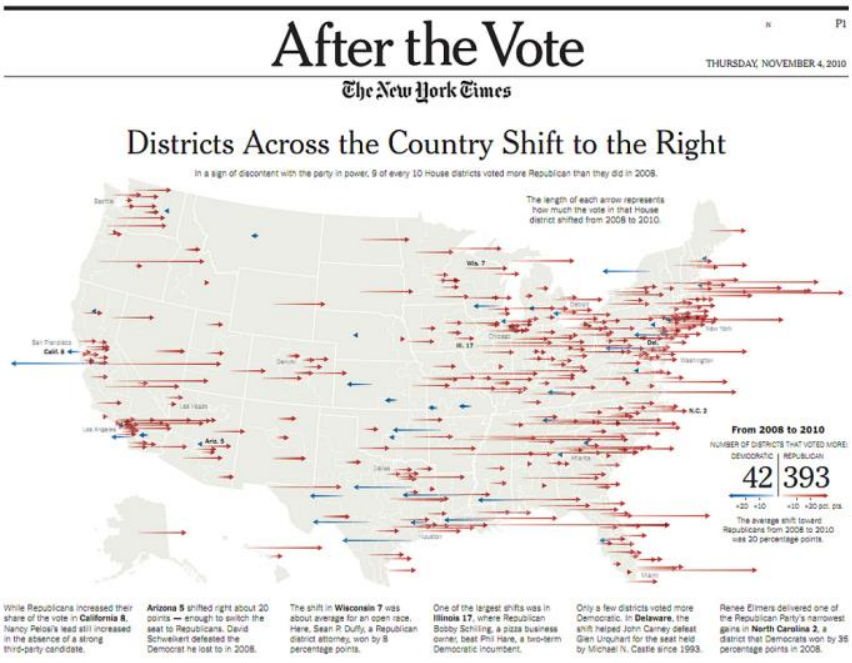
- They orient the entire view for clear readability.

The more you use the same techniques, the more polished your view will become.

Here are some examples:



Source: New York Times



Source: New York Times

Color Choices

Color can enhance or detract from a viz. Used effectively, you can turn a dashboard from mildly insightful to instantly impactful. The following are some best practices:

Do

- Use sequential or diverging color schemes to encode continuous ranges of numeric values.
    - Use stepped color rather than the completely continuous ranges as stepped are easier to perceive
    - Typically use 5 colors or less in a palette
        - 4 colors = quartiles (bottom 25%, 50%, 75%, and top 25%), 5 colors = quantiles (0-20%, 21-40%, 41-60%, 61-80%, 81-100% of the values)
        - Knowing what the stepped colors ranges are actually representing (i.e. quartiles) can help you quickly identify the top and bottom performers and know what range your middle data is falling into
- Use color to encode categorical (Dimension) variables.
- TRY: Use one and only one color encoding per dashboard.
    - Choose the key issue your dashboard is tackling

Don't

- Don't use the same color for two different variables
- Don't use the same continuous color scale for different magnitudes
    - Draws a relationship that may be incorrect
- Don't use an overwhelming amount of color on a dashboard

Guided Analytics

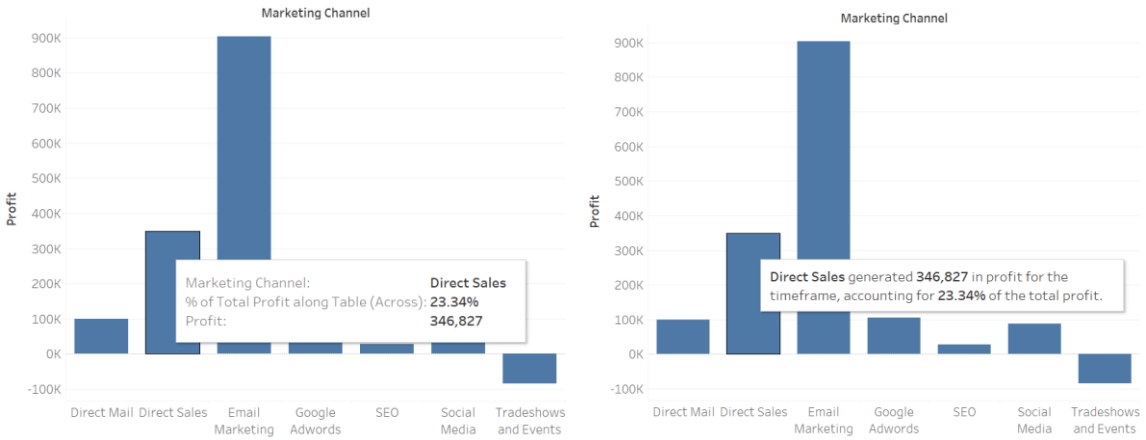
You want to avoid competing with yourself in a dashboard. Start with a high level insight and then help the user drill down to further levels of detail to understand further or know where to take action. This is called "Guided Analytics." This approach can best be utilized by using all three types of Dashboard Actions, in conjunction with the above principles.

- Guide their attention through the viz
- Answer one question at a time
- Allow them to drill down – hierarchies and pop-up charts

- Allow them to drill through – url actions passing key data as a filter
- Allow them to explore from different angles – parameters and story points
- Add annotations to highlight key points

Advanced Tooltips

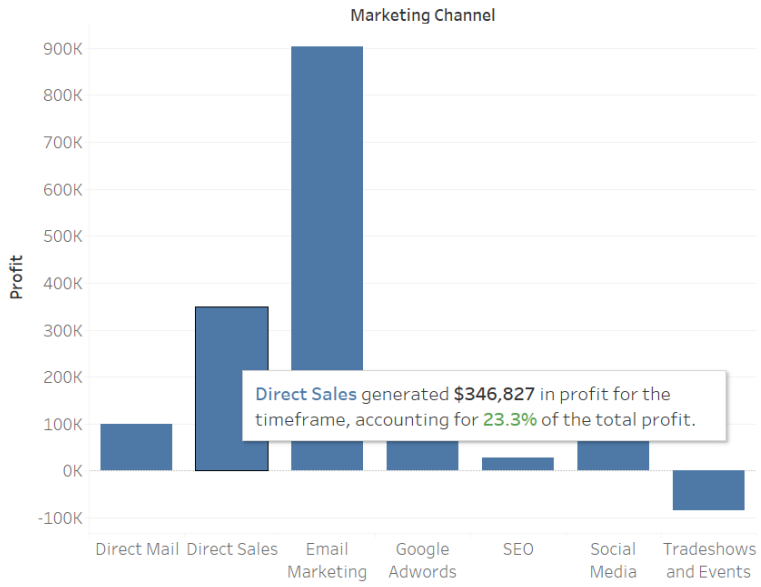
Tooltips provide more detail on demand, and if used effectively, can take a dashboard to the next level. This fits perfectly with the idea of guided analytics. Look at the difference between these two tooltips.



They both use the same information, but one uses natural language to make the insights more comprehensible. To create the second tooltip:

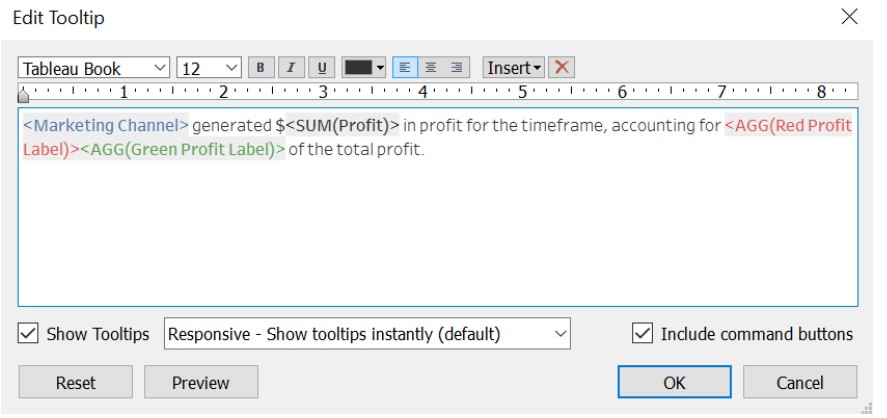
1. Drag Marketing Channel to Columns and Profit to Rows.
2. Drag another instance of Profit to the Tooltip shelf and use the down arrow to create a Quick Table Calc of "Percent of Total."
3. Click on the tooltip icon on the Marks card.
4. Delete the "Marketing Channel" header, leaving only the dynamic dimension field, denoted by angle brackets < >.
5. Delete the other two text fields in the tooltip, leaving only the dynamic dimension fields.
6. Now ensure your tooltip has the following sequence of text and fields: <Marketing Channel> generated <SUM(Profit)> in profit for the timeframe, accounting for <% of Total SUM(Profit)> of the total profit.
7. Bold only the three dynamic fields.

Now let's take this a step further, using a calculation to create some conditional formatting to produce the following tooltip:



1. Highlight the Marketing Channel dimension with your cursor and set it to the top blue color at the top right of the color strip on the drop down color menu. That color will be constant.
2. Double click the Percent of Total Pill that is on your Tooltip shelf. That will eliminate the pill's color and enable you to see the underlying Percent of Total calculation. Hit Ctrl A to select that whole formula and Ctrl C to copy it.
3. Now we're going to create 2 new calculations that we'll bring into our tooltip. Click the down arrow at the top left next to Dimensions, and choose "Create Calculated Field."
4. Type in the following formula: if sign(SUM([Profit]) / TOTAL(SUM([Profit]))) <0 then SUM([Profit]) / TOTAL(SUM([Profit])) else null end.
    1. Most of this formula is just pasting what we already copied – the percent of total formula.
    2. We start with an if statement, which as you recall has syntax If <condition> then <result> else <result> end.
    3. The sign function simply looks to see whether something is positive or negative.
    4. So, this formula is saying "if the percent of total profit is negative then show the percent of profit, otherwise (else) don't show anything at all."
5. Rename this calculation Red Profit Label.
6. Now copy this entire calculation and then click ok.
7. Create a new calculation, again by using your drop down arrow at the top near Dimensions or by right clicking in the Dimensions pane.

8. Paste the calculation you just copied in.
9. But, this time, change the "<" to a ">."
10. Rename this to Green Profit Label. Click ok.
11. Drag both the Red Profit Label and Green Profit Label to the Tooltips shelf.
12. Back in the Edit Tooltip window by clicking into it, we're going to change the last part of what we wrote by deleting the Percent of Total dynamic field and replacing it with both of the Label calculations we just created. It will look like this.



13. Color the Green Label with a green color and the Red Label with a red color to match the above image.
14. Click ok and mouse over the different bars in the chart. Notice how the color changes if it's negative or not. You now have conditionally formatted tooltips that provide a lot of additional insight and clarity.

Practice Exercises

Getting the formatting and aesthetics right often takes as much or more time as actually building the dashboard. Getting your dashboard to truly be a guided analytic, that conveys key insights at a glance, is hard work! That being said, you're becoming more pro in your skills, so it's time to embrace this challenge.

1. Create a new dashboard.
2. Start off and add a title: "Product Pricing Analytics."
3. Set the dashboard to Floating in the middle left side pane.
4. Drag out the "# Products by Price Point" chart we created. It was one of the later charts in our exercises.

5. Click on the chart. Now, click on the Layout tab at the top left. You'll see x and y and width and height coordinates. Set those to x = 5, y = 58, w =372, h =346.
6. Drag out the "Profitable vs. Unprofitable Products by Price Point" chart (the tab labeled Pricing Heatmap). It should also be floating.
7. Set the coordinates for this chart to x = 5, y = 412, w =372, h =356.
8. Eliminate the size legend and move the profit legend and size it to fit underneath the chart.
9. Click on the Product Category header and check Hide Field Labels for Columns.
10. Right click and Format the column headers of this chart. Set the font size to 8 and make them a lighter gray color.
11. Set the chart to be Fit Width.
12. Go back into the "# Products by Price Point" chart. Do this by clicking the square box and arrow icon at the top right.
13. Change the color scheme to Seattle Grays. This isn't the most important view, so we don't want to draw as much attention to it with the varied colors. Feel free to include a white or light gray Border in between the lines to separate them.
14. Now, go to the bottom and right click and create a new worksheet.
15. Drag out Order Quantity and Profit Margin to the Columns shelf.
16. Drag out Product Category and Product Sub-category to the Rows shelf.
17. Then drag out the Unit Price (bin) field, followed by Product Name and Avg Unit Price. (You'll have to set it to Average and make it Discrete by using the down arrow). All that is still on the Rows shelf.
18. Right click the Unit Price (bin) and uncheck Show Header. This will be important for our filter action later on.
19. On the Marks card, set the Order Quantity to a medium gray, again to provide detail, but not to call attention to itself.
20. On the Profit Margin shelf on the Marks card, drag another instance of Profit Margin to the color shelf. This is the key metric – we want to call attention to the margins.
21. Turn on labels on both metrics, from the Labels shelf.
22. Right click on the Avg. Price field > Format. Font = Dark Blue, Bold. Alignment = Center. This is the other key field – the price.
23. Unbold the headers at the top of the chart by right clicking > Format > font 8 and set them to the same gray color you chose for the previous chart.
24. Title this chart "Profit Margin per Product and Price Point."
25. Go back to the "Product Pricing Analytics" dashboard.

26. Drag this new sheet out. Also float. Set the coordinates to x = 395, y = 58, w =600, h =740.
27. Set the color legend for the "# Products by Price Point" chart to floating and drag it inside the first chart, bottom right corner. Right click on it and uncheck Show Header.
28. Drag the profit margin color legend to the top right above the right chart. You may need to click the headers in this chart and drag then down to provide enough white space to fit the legend above the chart and below the chart title.
29. Double click the Profit Margin label on the legend and set the font to 8, and unbold it. Then center it. We want this to be descriptive, but not call attention to itself. Click ok. Now shrink the width of that legend a bit.
30. Set each of the chart title fonts to be a dark blue. Do this by double clicking each one and editing the font.
31. Now unclick anything that is selected in the dashboard. You should now see at the bottom left a section for size. It will default to either Desktop or Automatic. Set it to Letter Landscape to provide a little more space.
32. Now add an action from the Dashboard menu. Action > Add Action > Filter.
33. Check the first box in the top pane. Set "Run Action On" to Select.
34. Uncheck the first box in the bottom pane. Set "Clearing the Selection will" to Exclude.
35. Click ok.
36. Now click the white space in the top left chart. The rest of your dashboard should go blank except for the titles.
37. Create a text box from the left hand side. Drag it out. Type in "If this is blank, click on a Price or bar segment in the top left chart. This will filter to the details you want." Align this towards the top of the blank space on the right hand side.
38. Now click the down arrow on this text box. Select Floating Order > Send to Back.
39. Now click again in the first chart and see what happens. You should quickly see which Product Categories and Sub-categories have problems and which Products have the worst Profit Margins. This should lead to instant insight and someone can take action on it.
    a. Notice how we kept the color schemes consistent, and minimized certain things in favor of the insight we wanted to convey.
    b. We laid out the charts in such a way that the user is guided through the charts step by step and the key points are emphasized.

c. We used descriptive titles and instructions hidden in the dashboard to help users engage with it.
d. And we used labels appropriately so that the end user could get a clear sense of the magnitude of the issue.

Custom Shapes & Custom Color Palettes

Tableau has several built in shapes and color palettes to add variety and meaning to a viz or dashboard. But sometimes, you want to add something additional to give extra meaning or visual appeal. Tableau gives you the ability to add additional shapes and additional colors to your palettes to achieve the look and feel that you want.
To add custom shapes, you'll typically need to do a google search to find some icons. I typically edit them in Powerpoint, but you can use whatever graphics program like you. Then you'll need to save each one individually as a .png file to your computer.

1. Now, go to My Documents > My Tableau Repository > Shapes.
2. Right click and create a new folder that's descriptive of your images (or whatever you want to call it). For example, the folder Product Subcategory could contain images relating to the subcategories.
3. Save each image as a .png file in that new folder you created.
4. Go back into Tableau. Set the drop down menu on the Marks card to Shape.
5. Click on the Shape icon that now appears on the card.
6. Click Reload Shapes.
7. Now, click the drop down menu and find the custom icons you imported. These are now available for you to use in a viz.

Here's an example of someone who has imported the logos of the different teams as well as a line chart showing their trends along with an embedded video from YouTube of their performance. Pretty sweet.

To create a custom color palette, first find the color palette that you want. I often will look at paint chip samples online or search up "bold color palette" or "neutral color palette" etc. to find what I'm looking for.

Once you've found an image of a palette you like, use the free Interworks color tool to get the codes you'll need for Tableau. http://powertoolsfortableau.com/tools/color-tool

Follow the prompts on their site to upload your image and get the codes. Once you have the codes:

1. Go back into My Documents > My Tableau Repository.
2. Look for the Preferences file or Preferences.tps.
3. Open this up with a text editor like WordPad or Notepad.
4. Paste the code that the Color Tool created into that file.
5. Save and close.
6. Go back to Tableau and open the Color mark and see your new palettes.
7. As long as you publish to Server, or save your workbook as a packaged workbook, these new colors will always stay with the workbook, regardless of whether your end user has gone and downloaded the same color palette, which is great.

Advanced Segmentation

We've touched on various components of how to take your analyses and insights skills to a higher level, including sorting and filtering data, to being able to see more details, or using more insightful charts and Quick Table Calculations, like percentages of total and percent differences. Another key ability of any data analyst is to be able to effectively segment the data such that actionable insights emerge.

Segmentation can be used to answer the following types of questions:

1. Who are our customers?
2. What do they buy?
3. What are their needs?
4. Where can we find these people? (based on location, actions, social activity, etc.)
5. What kinds of messaging, communication and advertising will resonate with these groups?
6. What are common patterns in behavior?
7. Where or from whom can I find similar products?

By dragging any Dimension onto Rows or Columns, we are creating a segment, such as Region or Marketing Channel. But we can become more analytically sophisticated in our segments through the use of Dynamic Sets and Clustering.

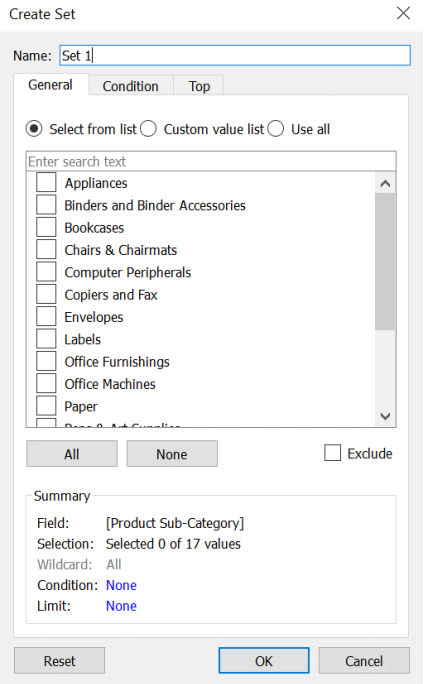
Dynamic Sets

Sets can either be static or dynamic. A static set is a manual group you create. For example, you could create a set called Continental USA and include all contiguous 48 states, excluding Alaska and Hawaii. Or you could lasso a group of products and group them by right clicking and choosing Create Set. These are just slightly more advanced than normal Groups, which we discussed much earlier in the book. Let's not spend any more time there.

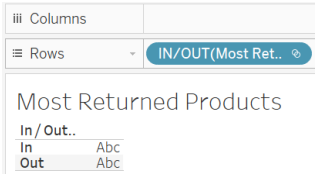
Top N Sets

We want to explore Dynamic Sets, which recalculate based on updates to the data. As an example, let's create a set called Most Returned Products, which will update with whatever products get returned the most.

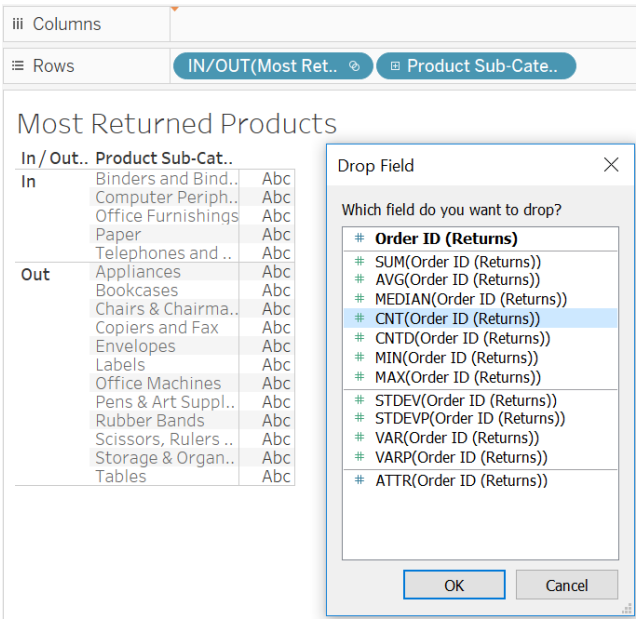
1. To start, you need to right click on the Dimension you want to create a set for. In this case, right click on Product Sub-Category, and then choose Create > Set.
2. A dialog box will appear with options: General, Condition, Top. Choose Top.



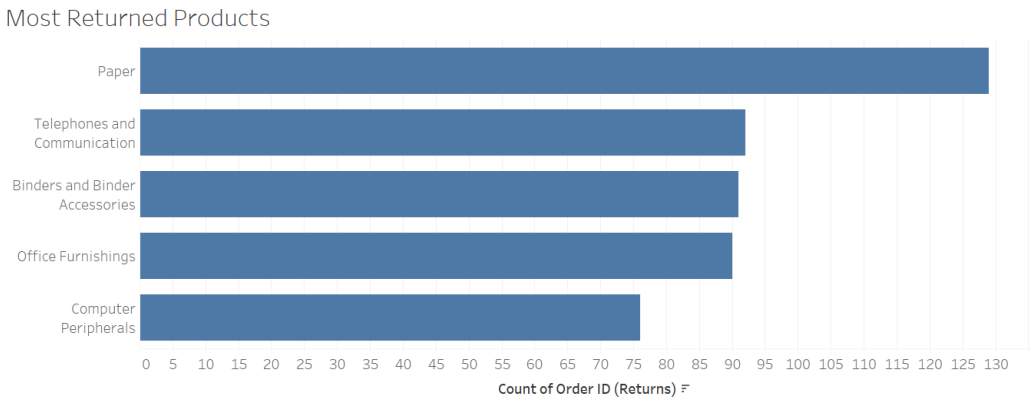
3. Click the By Field button. Change the number to 5.
4. From the drop down menu below that, choose Order ID (Returns).
5. Set the aggregation to the right of that to Count.
6. Rename this Set to Most Returned Products. Click ok.
    - The newly created Set will now appear in the bottom left corner.
7. Now we'll use this set in a viz. Open a new worksheet. Drag your set out to Rows. It will look like this image. We have the option to use the set in two ways, either showing those items within and out of the set, or to show just the members in the set.



8. For now, drag out Product Sub-Category to the right of your set on the Rows shelf.
9. Now, right click on Order ID (Returns) and drag that to Columns (while still holding down the right click key).
10. A dialog box will appear, overriding Tableau's defaults. Choose CNT for count.



11. This will create a bar chart that clearly shows its brought the most returned items into the "In" group and placed everything else in the "Out" group.
12. Now, click the drop down arrow on the In/Out set pill and choose "Show members in set." Remove Product Sub-Category from the Rows shelf.
13. Click the Sort Desc icon and we get the following viz:

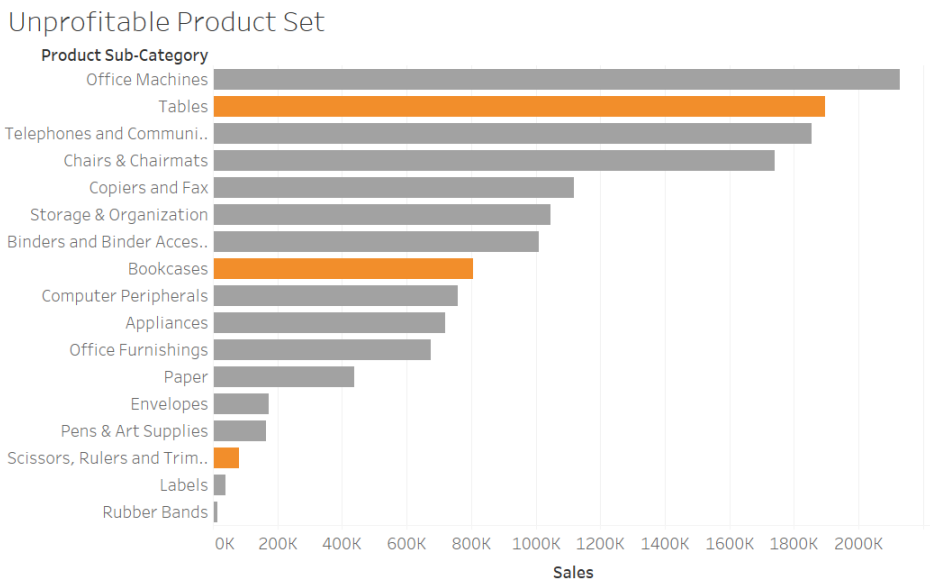


The great thing about this is it's a calculation, meaning what is "In" the set is dynamic based on the underlying values. If the store introduces a line of computer mice that don't work, they'll all likely get returned and we'll see that product appear on the list.

Conditional Sets

We can also create a condition that puts items into the Set if they are above or below a certain threshold, say Profit below $0. The previous Set we created was a Top 5. Using a condition grabs anything that meets the criteria you set.

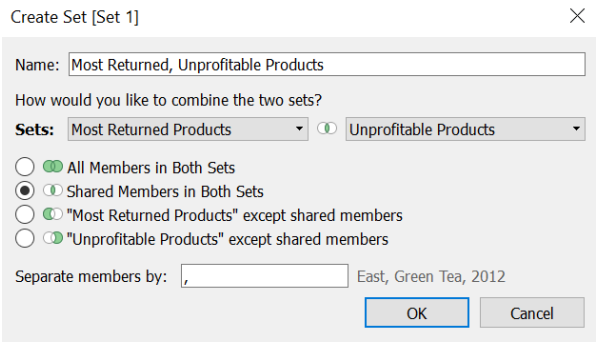
1. Let's create a new Set. Right click on Product Sub-Category again and choose Create > Set.
2. Click on the Condition tab.
3. Choose the button By Field.
4. Set that drop down to Profit and the aggregation to Sum. Set it to <= from the drop down and then type in "0" in the box.
5. Rename this set to Unprofitable products. Click ok.
    1. Your set will appear at the bottom left like the previous one we created.
6. Create a new worksheet and title it Unprofitable Product Set.
7. Drag Product Sub-Category to Rows. Drag Sales to Columns.
8. Sort Desc.
9. Drag your newly created Set to Color.
10. Click the drop down arrow on your color legend to change the color and choose orange for those that are "In" the Set. We can now quickly see that Tables are the 2nd highest selling item yet are unprofitable. We can quickly take action on this. Once we correct our unprofitable ways, Tables will drop out of the Set.



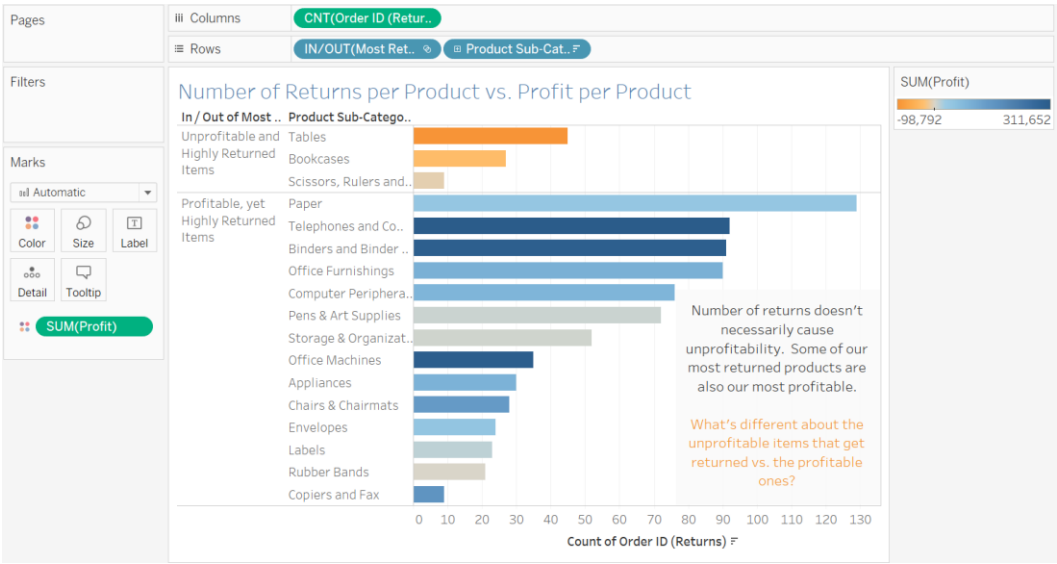
Combined Sets

Lastly, we can create a combined set, which outputs the results from the interaction of 2 sets, for an even more specific, yet still dynamic, grouping.

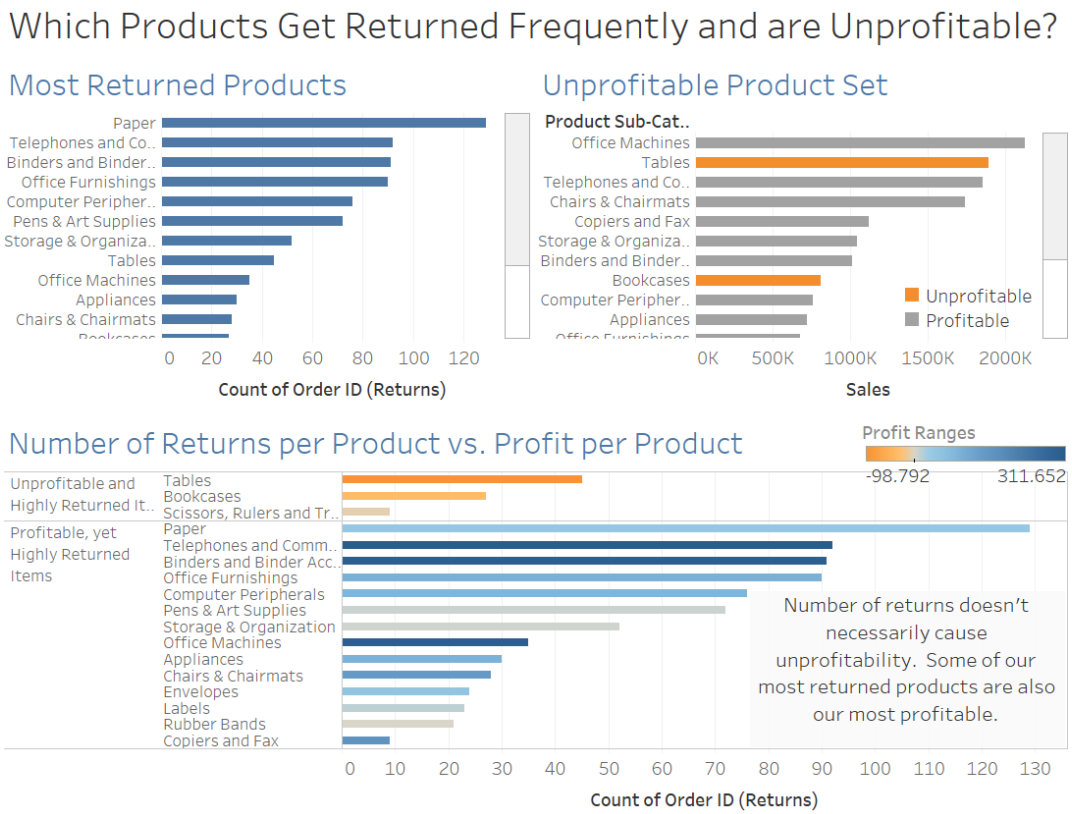
1. Right click on one of your sets at the bottom left. Choose Create Combined Set.
2. Name this Set, "Most Returned, Unprofitable Products."
3. Ensure one set has been selected from each of the drop down menus.
4. There are a variety of options we can choose from, but in our case we want "Shared Members in Both Sets." Choose ok.



5. Create a new worksheet. Drag your newest combined Set out to Rows.
6. Drag Product Sub-Category to the right on the Rows shelf.
7. Right click and drag Order ID (Returns) to Columns. Set it to Count. Notice how there is no "In" set; it's all Out.
8. Right click on your Most Returned Products Set at the bottom left and choose Edit Set.
9. Click the Top tab and change it from 5 to 20. Notice how the viz changes.
10. Now, drag Profit to Color.
11. Sort Desc.
12. Title the worksheet "Number of Returns per Product vs. Profit per Product."
13. Include an annotation by right clicking Annotate > Area.
14. Type in: "Number of returns doesn't necessarily cause unprofitability. Some of our most returned products are also our most profitable. What's different about the unprofitable items that get returned vs. the profitable ones?"
15. Finally, right click on the In/Out labels on your chart. Choose Edit Alias.
16. For "In," change that to "Unprofitable and Highly Returned Items."
17. For "Out," change that to "Profitable, yet Highly Returned Items."



We could tie those three separate charts into a dashboard, with a little additional formatting, like so:



Practice Exercises

The VP of Sales wants to know the products that are purchased the most compared to products that are unprofitable. She doesn't want to incentivize her sales team to hit sales targets that end up not being profitable for the company. She asks you to help her have that insight available at her fingertips, no manual analysis on her part required nor time to wait for it to get updated by an analyst somewhere, and it needs to be dynamic.

1. Right click on Product Sub-Category, Create>Set.
2. Choose the Top tab.
3. Set By Field to Top. Then, in the next box to the right, click the drop down, and choose *Create a new parameter*.
4. Title that parameter Top N.
5. Ensure the data type is set to Integer and the Allowable Values are set to Range.
6. Set the Maximum to 10 and click Ok.
7. Then, back in the Sets editor, choose Order ID from the large drop down menu and ensure the aggregation next to it is set to Count.
8. Title this Set "Most Ordered Products".
9. Drag this Set onto a new worksheet on Rows. Then, drag Product Sub-Category to the right of that pill, also on Rows.
10. Right click on Order ID and drag it to columns. Set the aggregation to Count when the aggregation window pops up.
11. It should be clear what the most ordered products are, and you can determine how many products make it into those groupings by adjusting your Top N parameter, which should automatically appear up at the top right. If it doesn't, go down to the Parameters section at the bottom left, right click on it and Show Parameter Control.
12. Now, create another set again on Product Sub-Category (right click the dimension and go from there).
13. This time, go to the Condition tab.
14. Set the By field to Profit, the aggregation to Sum, and then set it to < 0.
15. Name this Set "Unprofitable Product Lines".
16. Drag this on Rows in between the 2 pills that are already there.
17. Rename this viz to "Most Ordered, Profitable vs. Unprofitable Products".
18. Set your Top N filter to 10 and ensure that the Profitable set appears before the Unprofitable ones. If they aren't in that order, click on the Unprofitable label in the viz and drag it down.

19. Change the labels on your first set to be more descriptive than In/Out. Right click on the one that says "In" and Edit Alias to "Most Ordered". Change "Out" to be "Less Popular".
20. Drag this most recent set also to the Color shelf.
21. Save the workbook (probably a good time to do that!)

Advanced Analytics

Part of Tableau's greatness is its ability to make advanced analytics accessible within a few clicks, as we saw with clustering and Quick Table Calcs. This, in turn, can make you a better, more insightful analyst. What otherwise would have required complicated Excel formulas or coding knowledge, can be done in a few clicks. The trick is knowing what's available to you, when to use it and how to interpret it or explain it to others. This section on Advanced Analytics helps you do that.

Reference Bands and Reference Distributions Explained

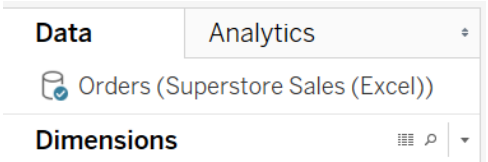
We've already examined how adding a reference line to a chart can provide a lot of insight, highlighting issues like:

- Are we above or below a desired threshold?
- What's the average or median of all these values?
- Where are the outliers and what should I be paying attention to?

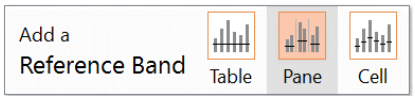
Reference bands and distributions help you take that a step further and see how data falls within certain ranges, rather than just one point.

Reference bands

1. Create a new tab and drag Sales to Rows and Product Sub-Category to Columns.
2. Now click on the Analytics tab at the top left.

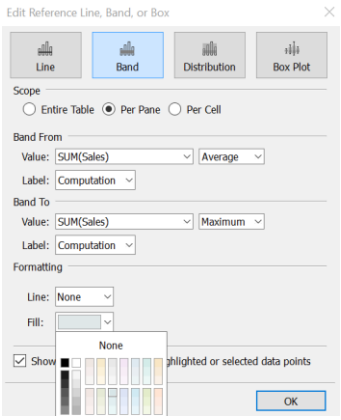


3. Drag the Reference Band option onto the chart. A window will pop up asking if you want it on the Table, Pane, or Cell. Put it on the Pane option.

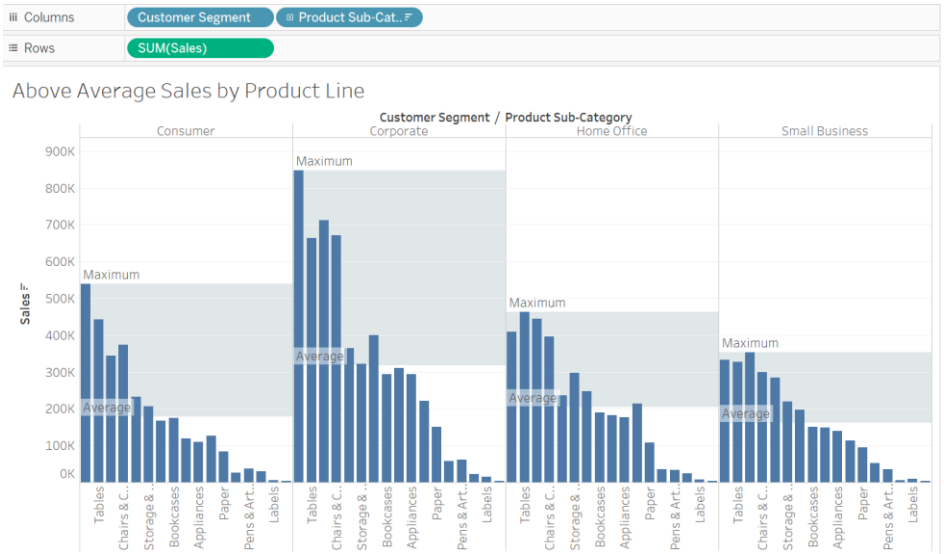


1. The cell option applies to each cell or data point in your chart (i.e. each bar, or each dot). The pane applies to whatever is in the pane, denoted by gray line dividers. The table is the entire table. For a view with only one dimension, the Table and Pane will result in the same. But, we'll add more soon that will make the Pane the more useful choice.

4. The reference band window will appear. Set the Band From value to be the Average Sales Amount (i.e. change it from Minimum) and leave the Band to at Maximum.
5. Set the Fill color to be the top green color on the bottom color strip. Click ok.



6. Sort the chart in desc order.
7. Now, drag Customer Segment to Columns in front of Product Sub-Category.
8. Title the chart "Above Average Sales by Product Line."



9. Review the chart for insights that stand out, like the Small Business segment's sales per product line are closer together, meaning we can expect all sales to that segment to be more uniform and equal. Corporate clients, on the other hand, tend to be higher than the other segments in almost every category, but have a much wider span between average and max sales, meaning more variability in purchase behaviors.

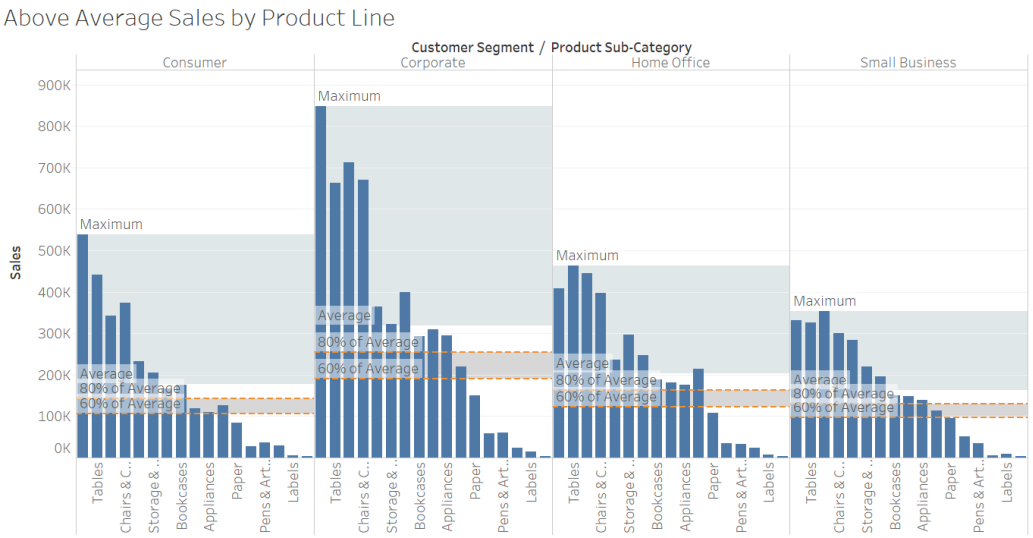
Note: these can be combined with parameters to make them more dynamic and flexible and allow you to control the ranges.

Reference Distributions

Reference distributions go even further, allowing you to encode multiple values in ranges. We'll look at each one using the same chart we just built.

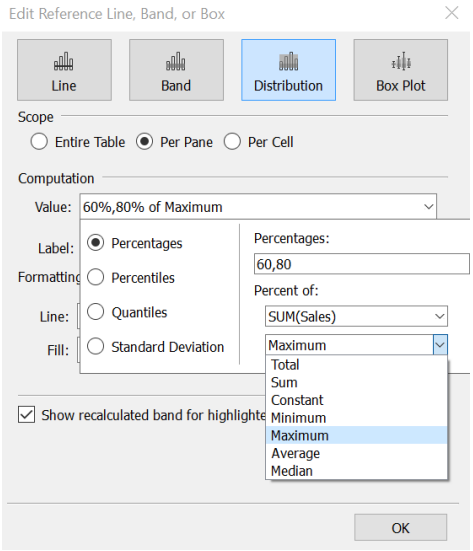
60, 80% of Total, Percentiles

1. On the Analytics pane on the left, click Distribution Band and drag it out on top of the chart. Again, be sure to select Pane from the window that pops up.
2. The Reference Distribution box will appear and everything should be defaulted. Just review it to ensure it's set to Per Pane, and the value is 60, 80% of average.
3. On the Line drop down, make it a dashed line and changed the color to orange.
4. Uncheck the box that says "Show recalculated band for highlighted data or selected data points."
5. Click ok.



This chart is showing us the range from Avg. to Max value per Customer segment (created in our previous section) and now it's also showing us those product lines that are 60 and 80% of the avg. respectively. But we don't really care which lines are 60% of the average; we already know those are performing below average. We'd rather see which products are within 60 and 80% of the maximums. With a little focus and effort on those products, we could bring them up closer to our top selling products.

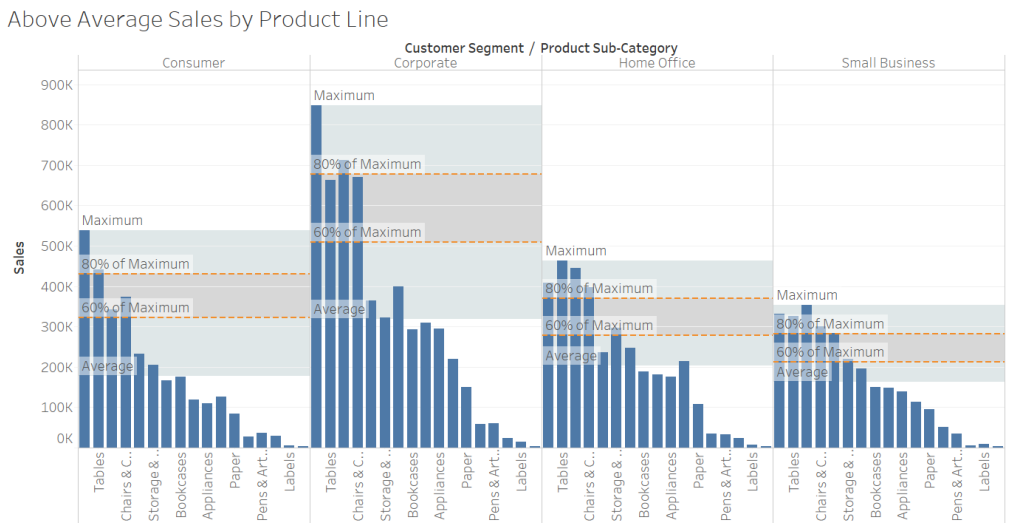
6. Right click on one of the orange lines. Choose Edit.
7. Change the Value to 60, 80% of Maximum by clicking the drop down menu and changing it from Average to Maximum.



8. Uncheck the box that says “Show recalculated band for highlighted data or selected data points.”

We can now see which products are close to the top performers. This is most useful when looking at goal or quota attainment. We can then see who’s or what is average, what’s below average, and the breakout of our top performers.

Note: You can also change the percentages from 60 and 80 to whatever you want. Maybe you want to see the top and bottom 10%, so you could change it from 60, 80 to 10, 90. Also, while everything there used Percentages, you can also use Percentiles, which work mostly the same way.

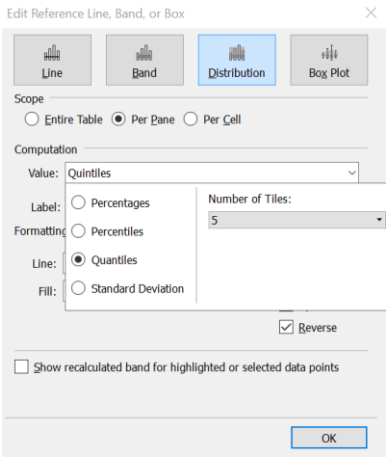


Quantiles

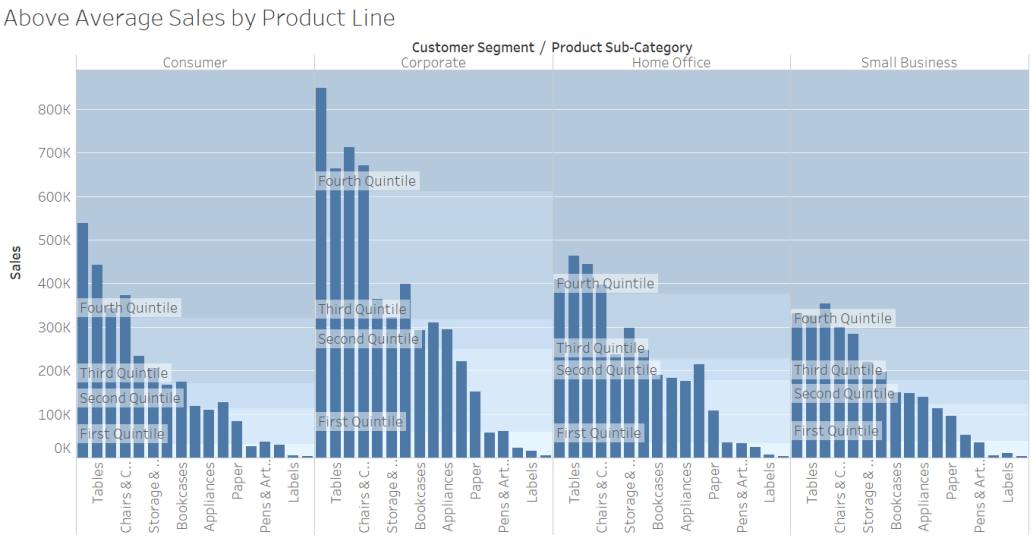
A quantile is like a percentage or percentile, but typically a larger segment. Quantiles are five segments – 20%, 40%, 60%, 80%, and 100%. Quartiles are quarters (obviously in increments of 25%), terciles are thirds (increments of 33%), and a decile is 10 cuts, with each segment equaling 10%. I normally leave it at either quartiles (4) or quantiles (5). Note that these segments aren’t necessarily equal; they’re dependent on the spread of your data. Maybe most of the values fall at the bottom 20%. The color ranges on your chart will reflect this.

1. Right click on the sheet we were just working on and duplicate it.

2. Click the reference band lines on the chart and while holding down your click, drag them off the viz.
3. Click the Analytics tab at the top left and then drag out the Distribution Band, set to Per Pane.
4. Set the Value to Quintiles from the drop down menu.
5. Change the Number of Tiles from that drop down menu from 4 to 5.
6. Check the Fill Above and Fill Below box on the right side.
7. Change the Fill color from the drop down menu to Blue Light.
8. Finally, check the Reverse box (seen in the image above). Click ok.



We now have a color banded chart, showing the values that fall in the bottom 20% on up.

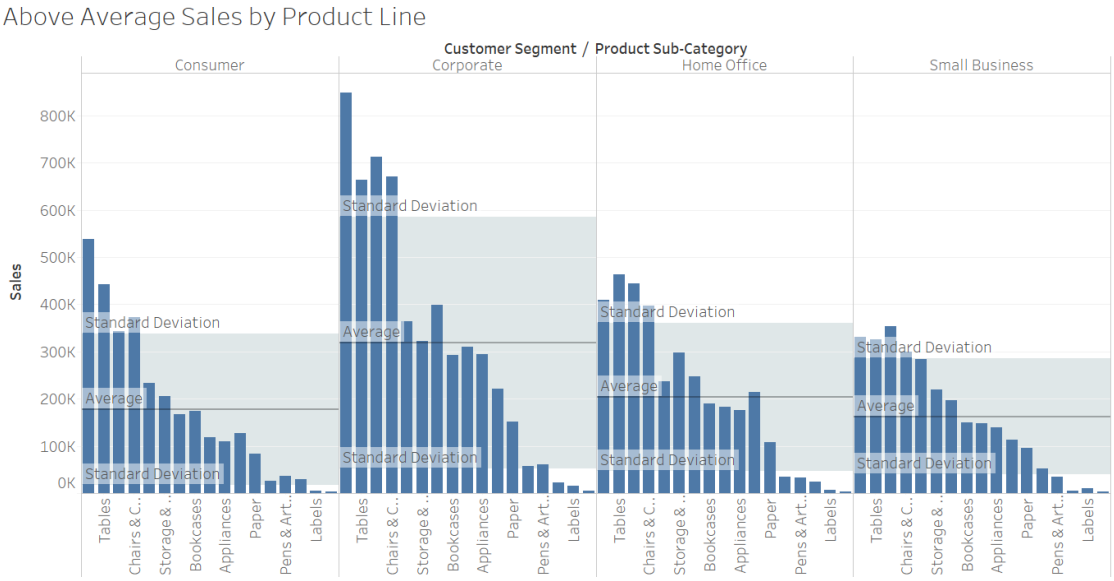


Standard Deviation

The standard deviation is a measure that is used to quantify the amount of variation or dispersion within a set of data values. A low standard deviation indicates that the data points tend to be close to the average (also called the expected value) of the set, while a high standard deviation indicates that the data points are spread out over a wider range of values. Note: Standard deviation, like average, is affected by outliers and is not the best measure of variation for skewed distributions.

We can use these distributions in Tableau to visually show the spread of our data around the mean or average.

1. Right click and duplicate the chart we've been working on.
2. Click on the reference distribution lines and while holding down your click, drag the distribution off the viz.
3. Click on the Analytics tab at the top left (this sequence should be starting to feel familiar by now ☺) and drag the Average line out onto your chart, setting it to Per Pane.
4. Go back to the Analytics tab and now drag the Distribution Band out to the viz.
5. Change it from the default 60%, 80% to Standard Deviation. It will default to -1,1. Leave it at that.

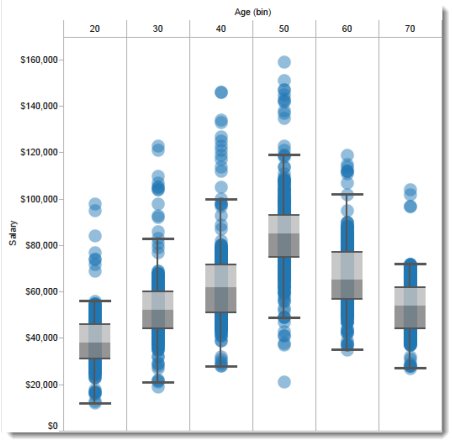


We've now got a view showing 1 standard deviation above and below the mean or average. These are fairly wide bands, indicating our data is pretty widely distributed.

While each of these kinds of reference bands and distributions won't make it into every chart or dashboard that you create, it's useful to know how to use them when exploring your data and understanding how it's behaving, as well as being able to visually showcase these insights to others.

Box Plots

A box plot looks like this:



It allows you to compare a range of values across several segments. Let's say we wanted to see the breakout of ages for our employees. We could use a histogram or bar chart to show how many people fall into each of the age buckets. If we wanted to stratify salary ranges, we could take the same approach. But, what if we want to see salary ranges per age range? That becomes a much harder problem to visualize.

Box plots are a simple way of accomplishing that. They show ranges of data, or distributions, across one or multiple segments. In the chart above, we can see the distribution of salaries of people in their 20's (the first column in the chart). We can see that the average salary is just shy of $40,000, but that we have some outliers at $100,000 and $15,000. The $100k per year 22 year old might be a data scientist, fresh out of school with a dual major in statistics and business. The 22 year old making $15,000 probably didn't go to college and is working more of a minimum wage type job.

How to Interpret Box Plots

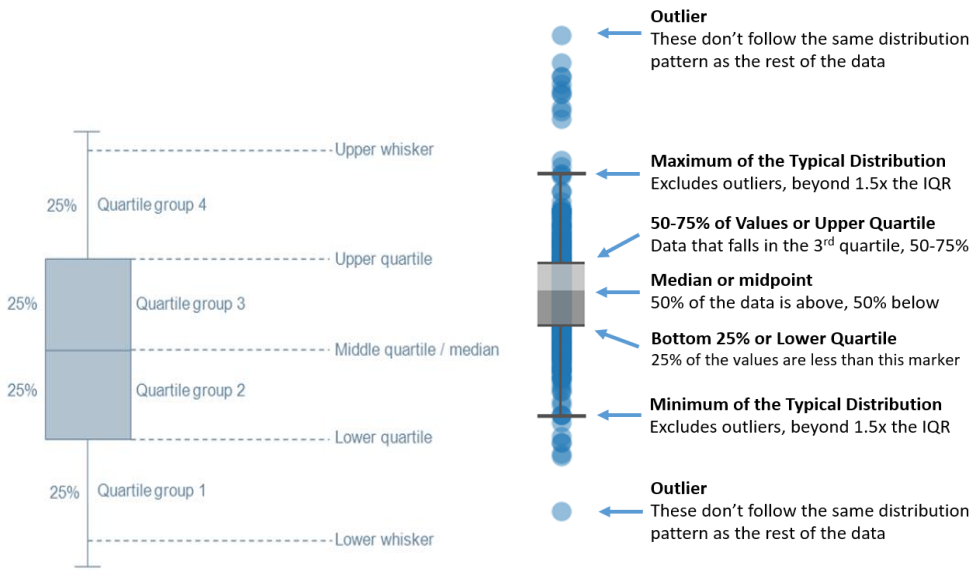
The line in the middle of the shaded box, or the dividing point between the two colors, is the median or midpoint of all the data values in the range.

The shaded area on each set of dots contains the middle 50% of all the data. This shaded area is known as the Interquartile Range. The bottom 25% of the data is below the shaded box. The top 25% is above the shaded box.

The whiskers typically represent 1.5 times the interquartile range (AKA the IQR, which is the length of the shaded box). So, the bottom whisker extends to the minimum point no further than 1.5x the IQR, and the top whisker extends no further than 1.5x the IQR.

The points at the very end represent outliers, if there are any. These are points that extend beyond the rest of the distribution. They are definitely worth investigating, and can often be the most useful pieces of data in your set. They indicate some anomaly in

the data, like a data error, or they indicate examples where the normal pattern breaks down for a good reason, and understanding why can lead to major new insights. If you decide they are skewing your data too much, you can exclude them to focus on the otherwise normal patterns.



Building a Boxplot

1. Drag Region to Columns and Profit to Rows. This will build us a bar chart by default.
2. Drag Customer Segment onto Columns, to the right of Region.
3. Click on Show Me and choose the Box Plot option.

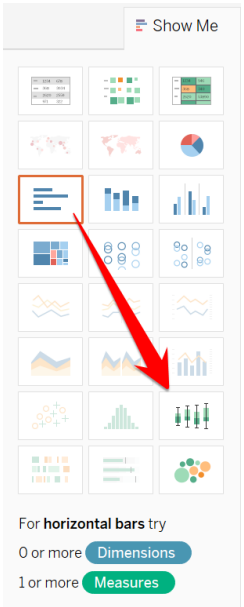
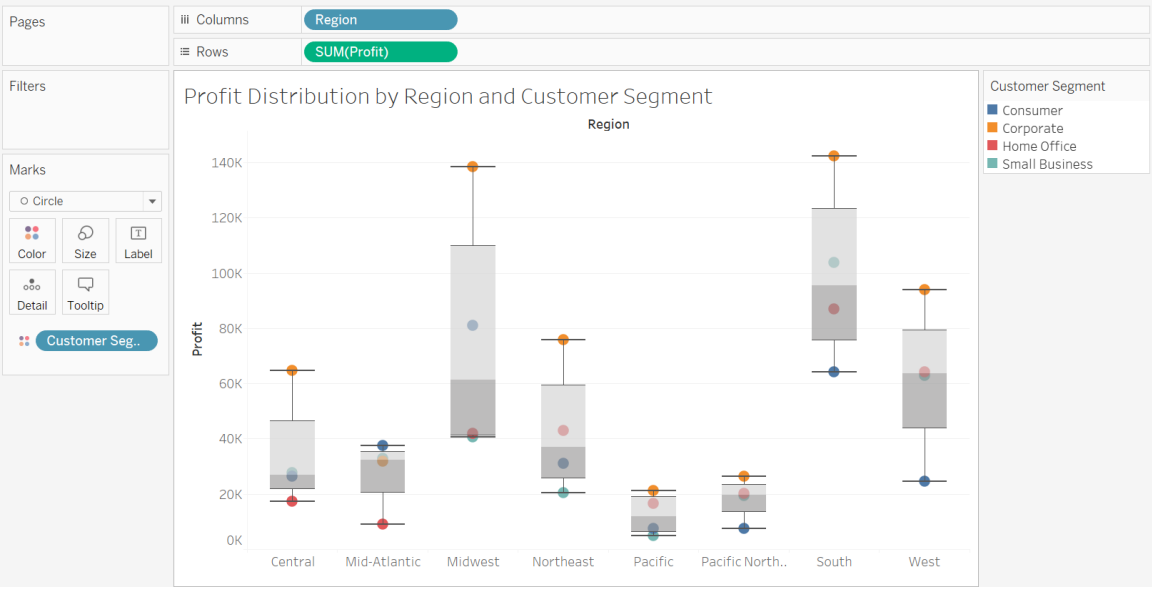


Tableau will automatically build you a box plot. However, it may not be organizing the pills in the order you want. In our case, we want Region on Columns and Customer Segment on Detail.

4. Click on the Analysis menu at the top.
5. Click on "Cycle Fields" towards the bottom of the menu. This should create the right box plot structure, with Regions as our headers running along the bottom, and each circle representing a Customer Segment.
6. Now, move the Customer Segment pill from Detail to the Color shelf, by dragging it in.
7. Set the view to Fit Width from the drop down menu above the chart.

8. Title your chart “Profit Distribution by Region and Customer Segment.”
9. Lastly, let’s size up our dots just a bit, by clicking on the Size shelf and moving the slider up somewhat.



We can now very quickly see that in most cases, Corporate customers are the outliers. They are our most profitable. Conversely, Consumers are our least profitable. Except in the Mid-Atlantic region, where Consumers are our top. This might be worth further investigation.

Forecasting

Tableau has some amazing built in forecasting features. This will get used in some domains more than others, but if it applies to you, it makes building a simple forecast easy. With each version of Tableau, the forecasting features get more advanced.

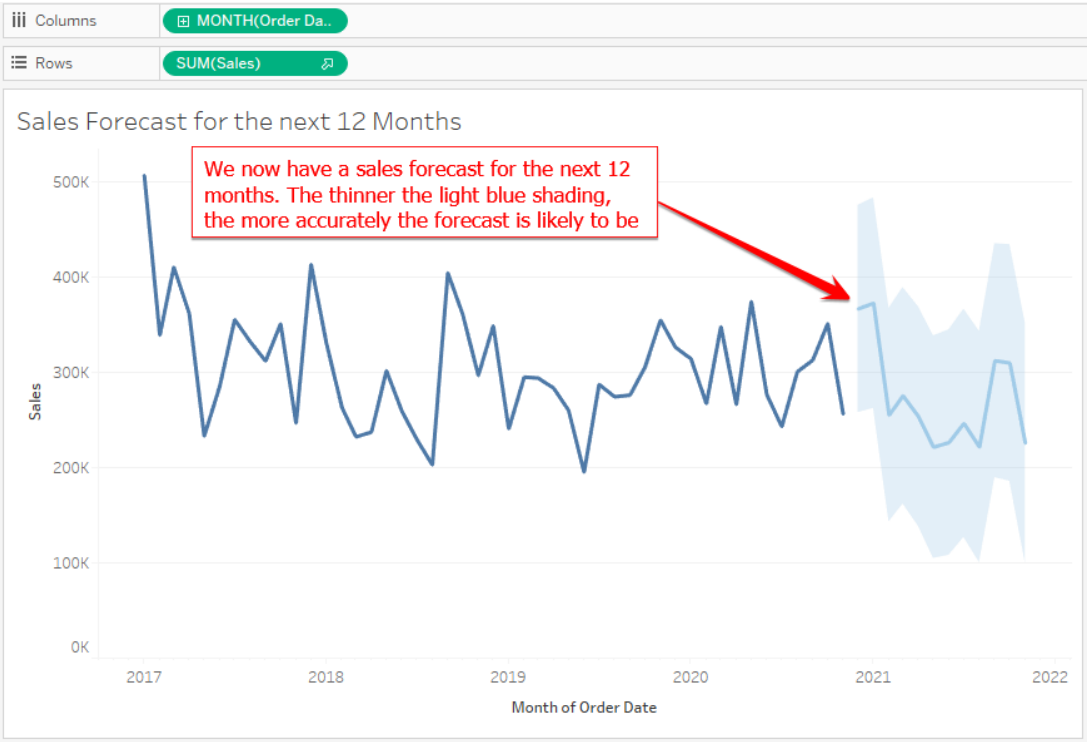
When are Forecasts Useful?

- Predicting sales trends and understanding spikes in demand
- Understanding seasonality
- Understanding how much inventory you need on hand at any given point in time
- For production and manufacturing planning on building and launching new products
- Predicting email response rates (though this would be even better if used in conjunction with Clustering)
- Understanding peak wait times in stores or lines, or peak usage times for mobile devices (to help in tuning cell towers)

- Predicting busy times at a hospital or call center by hour, day, or month to optimize staffing
- Essentially, any time you want to have a sense of what is coming in the near future

How to Build a Forecast in Tableau

1. On a new sheet, drag Order Date to Columns and Sales to Rows.
2. Change the Order Date to continuous by selecting the Month value in the second segment of date options. (i.e. Month May 2015)
3. Right click on the chart and choose Forecast.
4. You'll now have a sales forecast for the next 12 months. It gave us a straight line, so let's edit the forecast.
5. You can edit the forecast by right clicking on it and choosing Forecast Options. Under Forecast Model choose "Custom." Then we need to set Trend and Season to Multiplicative. Editing forecast options can get pretty technical, so do your research before showing the forecast to the CEO, just in case he throws you a curveball question.
    - For more information on when to use Additive vs. Multiplicative, see here: http://support.minitab.com/en-us/minitab/17/topic-library/modeling-statistics/time-series/time-series-models/additive-and-multiplicative-models/
    - But for the most part, you can be pretty comfortable using either. In this case, the multiplicative model was a little more accurate.
    - A good way to test this is to use some data to create a forecast and then test it against data you've withheld. For example, using data from 2017, 2018, 2019 and 2020, build a forecast for each month of 2021. Compare the values you that the forecast provides vs. the actuals for the year. They should be fairly close. You can test this with additive models or multiplicative, or some combination to see which one gets you closest overall. Know that the further out you project, the more variation you will get.
    - The great news is that Tableau will update the forecast each month, so you'll always be pretty accurate a few months into the future.
    - You can check accuracy with the formula (Forecast – Actual)/Forecast to determine percent variance.

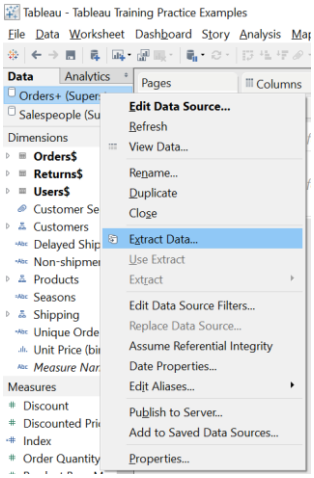


Advanced Extracts

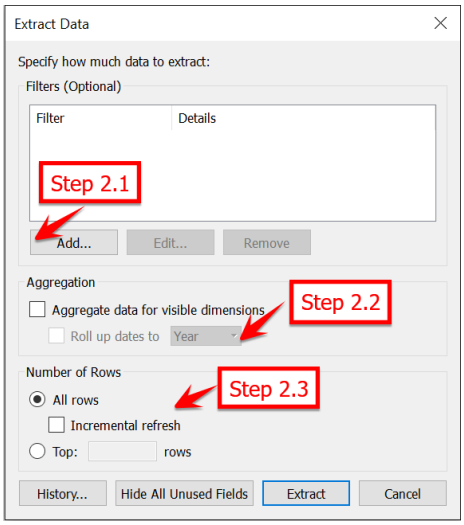
The last few sections of the book shift away from dashboard best practices and visualization types to ensuring the performance and accessibility of your dashboards are optimal. First, let's analyze Extracts. We discussed what they were previously. Basically, they build an in-memory database of your data so that you can query it much faster. You don't have to ping your actual database each time and re-render all the data.

As a general rule of thumb, use extracts always, because of the speed gains you will achieve. To create these:

1. You have two options. You can create it from the main data connection window via the radio button at the top right. Or, once inside a file, you can right click on your data connection at the top left and choose "Extract Data."



2. Upon clicking this, a new window will pop up.
    - The top portion will allow you to filter out any data that you don't want to come into your extract. Say all of your analysis will be aggregated at the region level, so you don't need to bring in states, as an example.
    - Speaking of, you can set the aggregation. An example might be you have data at the hourly level, but don't need it that granularly. You can aggregate it at the day level. When Tableau builds your extract, it will perform that process as part of the load. This will make the workbook faster when you access it.
    - Finally, you can tell it to rebuild the extract (i.e. database) from scratch each time you refresh it. This is a full refresh. You can also tell it to do an incremental refresh, only adding new data from the last refresh. This is useful as it speeds up extract creation time, but you have to be careful about old data changing. If a change has been made to data that is already in your extract, it won't capture that.



3. As a best practice, if you are creating workbooks that will utilize the same data, create an extract of the data and save it Tableau Server as a data source. Set it to extract and use Server to set a schedule. ***This automates your data collection.*** It's awesome. Every time new data comes into the database, your extract will go get it on the schedule and automatically update your workbooks with the new data.

Using Tableau Server

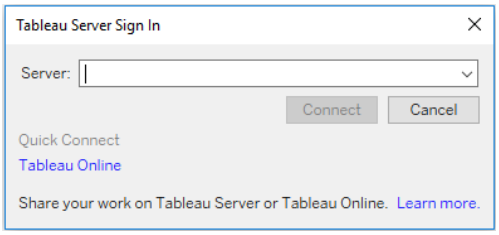
As described previously in the course, there are a few different ways to share data:

- Publish to PDF
- Export or copy an image into PowerPoint or an email
- Export the data to Excel (try to get people to not do this)
- Publish to Tableau Public (not recommended for sensitive, confidential, or proprietary data)
- Email a packaged workbook and have people access and interact with it via Tableau Reader (a great place to start if you don't have Server or someone doesn't yet have Server access)
- Publish it to either Tableau Server (the on-premise version) or Tableau Online (the cloud version)

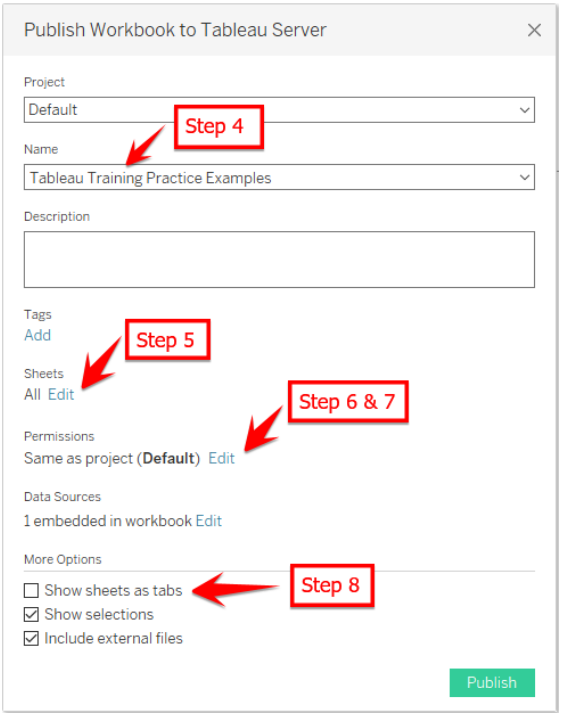
For many companies, they use Server or Tableau Online. These are great products for publishing an interactive dashboard to the web, while maintaining the security protocols your data demands. Your end user can interact with your published dashboards any time they want. And if you are using extracts that are on refresh schedules, your data will automatically update at the frequency you specify, keeping your audience always up to date.

To Publish a Workbook to Tableau Server:

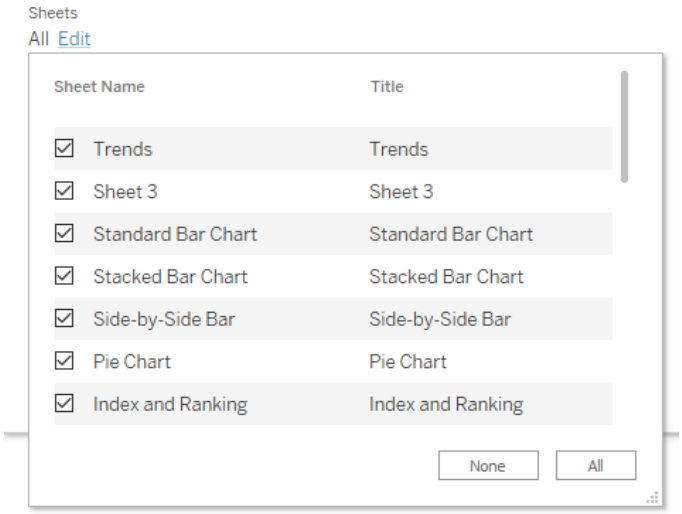
1. Go to the Server menu in the top navigation.
2. Click Publish Workbook.
3. Enter your authentication credentials (typically a single sign on username and password, or database validation credentials).



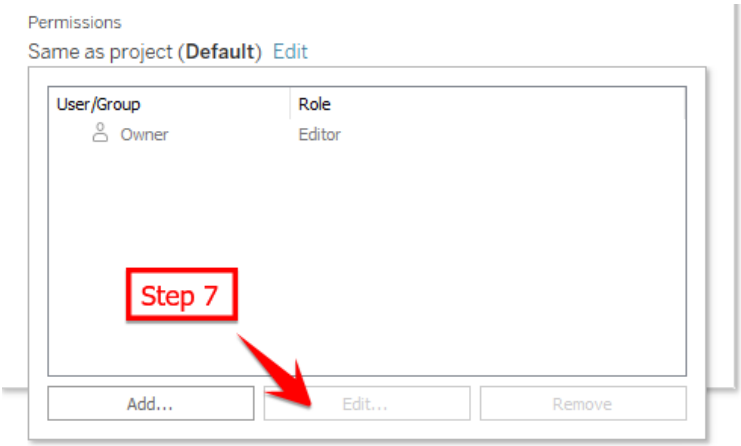
4. From there, make any changes needed to the default workbook name.



5. Under Sheets, click edit to select which sheets you want to publish.
    1. One best practice here is to hide the worksheets used in your dashboards, and only show the dashboards that you are going to publish. This keeps your file cleaner and eliminates check boxes you have to click.



6. Set any user permissions by clicking Edit next to Permissions. You can leave these as the default, or you can edit their permissions. There is a wizard with a series of radio buttons you can check or uncheck.
7. Finally, be sure to set your Authentication and Permission by ensuring that your password is set to "Embedded" not "Prompt User." You want consumption of your dashboards to be seamless, not involve extra steps.



8. Lastly, if you are using URL actions in your workbook, be sure to check the box in the bottom left that says "Show sheets as tabs"; otherwise, they won't work. If you aren't using URL actions, you can check them or uncheck them; it's up to you.

A New Perspective on Using Server

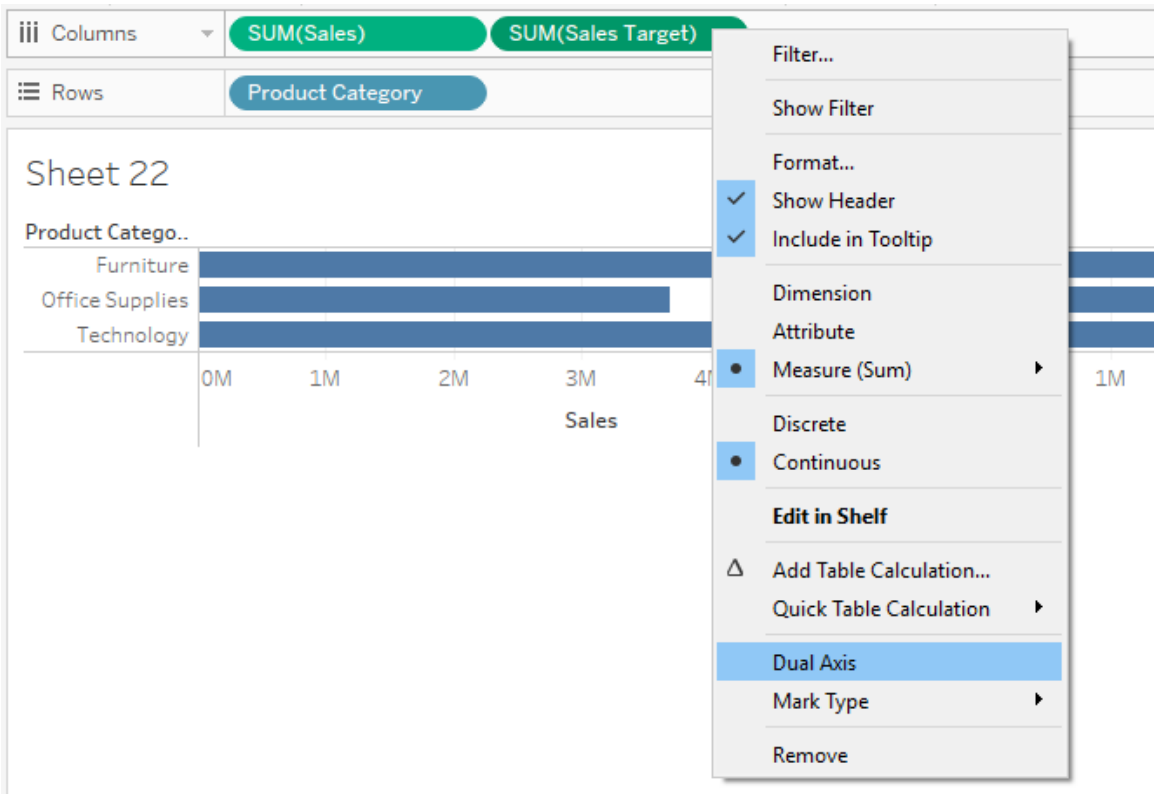
Too often, we take a snapshot of our data and paste it into PowerPoint, or people just want a data dump to Excel. It's an uphill battle to change this mindset, but completely worth it. Instead, if you can bring a published workbook to Server and interact with your data live in a meeting, that is much more impactful. As they then have further questions, you can drill down or drill in and see those details on demand, maintaining things on one view for easier comparison. This speeds decision making, and changes the discussion dramatically, and increasing participation. I would encourage you to shoot for this type of usage in your meetings. Also, as your people on the front line need data to make decisions, they have it instantly accessible via Server and can quickly get what they need and then take action, rather than putting in a request to you or to someone else, and waiting for a response to come back, which by then will have data that has changed from their initial need. When you achieve this, you've transformed a culture and can now truly compete on analytics.

Advanced Chart Types

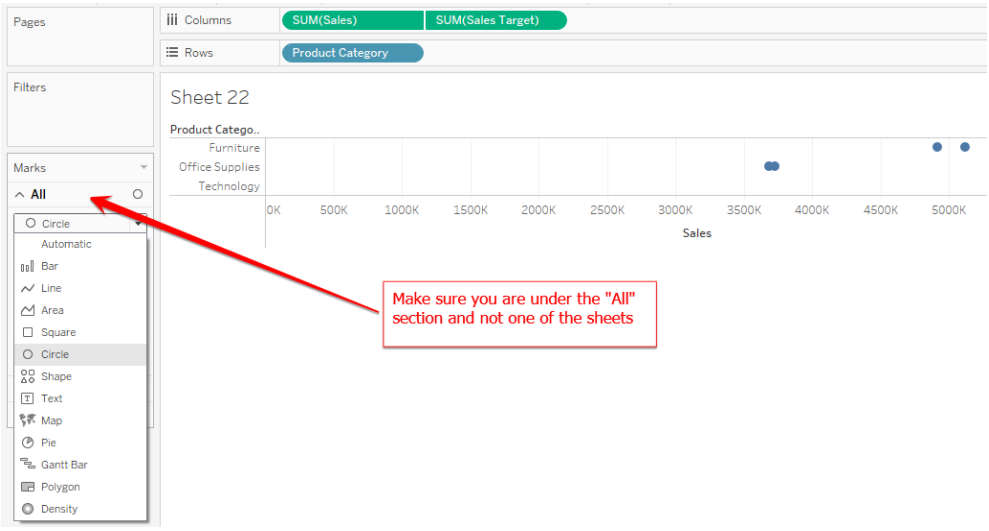
Bar in Bar Chart

A bar in bar chart plots two bars in the same space, one thicker, one thinner. It's great for comparing two measures or comparing one measure against a target. To create them in Tableau:

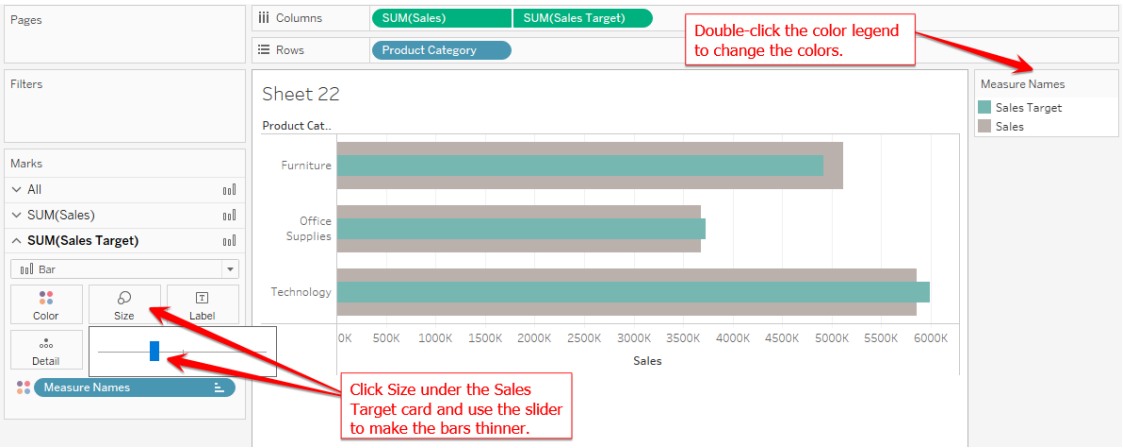
1. Bring out a dimension and two measures on a new worksheet. Let's make it Product Category and Sales and Sales Target.
2. Drag Product Category to the Rows shelf. Drag your Measures to the Columns shelf.
3. Right click on the down arrow on Sales Target and make it a dual axis.



4. Now, right click on the secondary axis (the one along the top of the chart) and select "Synchronize Axis." Now, right click again and hide the secondary axis by unchecking "Show Header."
5. Change the circles to bars under the All section of the marks card.



6. You should now have bars that overlap. To get the bar in bar effect, choose the measure that you want to show as the thinner bar – in this case, Profit – and click on the Marks card where it says Sum(Sales Target).
7. Click on the icon for Size and slide the bar down. The inner bar will thin and now shows you how Sales Target compares to sales.
8. Edit the colors to be what you want.



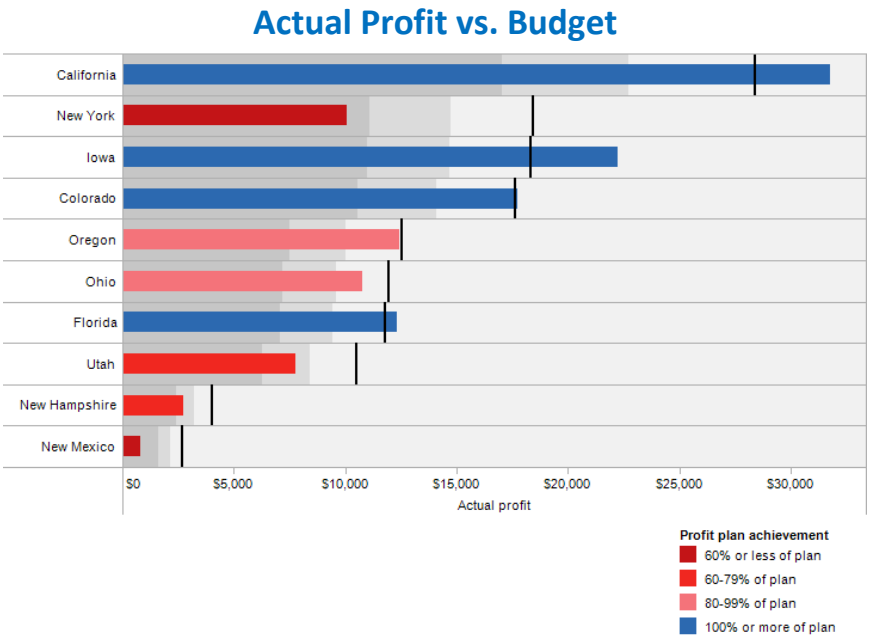
Bullet graphs

Bullet graphs take the bar in bar chart a step further. They are an excellent way to show progress towards a target as well as thresholds or ranges of progress or acceptability or tolerance. For example, if you wanted to see progress towards a goal and show anything below 60% of the way there in red, anything below 80% in yellow, and anything above in green with the target line at 90%, a bullet graph can do that. To create, there are a few possible ways.

The easiest is to use Show Me. Select 1 or more Dimensions and at least 2 Measures and click the Bullet graph option. Rearrange your measures and dimensions to get the desired view. You can also:

1. Follow the steps above to create a bar in bar chart.
2. Right click on the axis or use the Analytics pane to create a reference line.
3. Set the line to be Per Cell.
4. Set the line to be one of the metrics on your viz. If you don't want a particular metric on the view but want to use it as a reference line, put it on the Details shelf. Or create a custom parameter, allowing you or the end user to set the values.

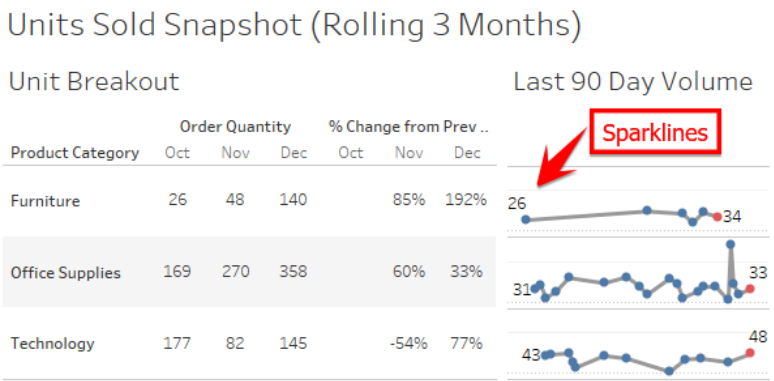
This is an example of what these can look like. The color is a calculated field using if/then logic to create the groups and is then dropped on the Color shelf.



Sparklines

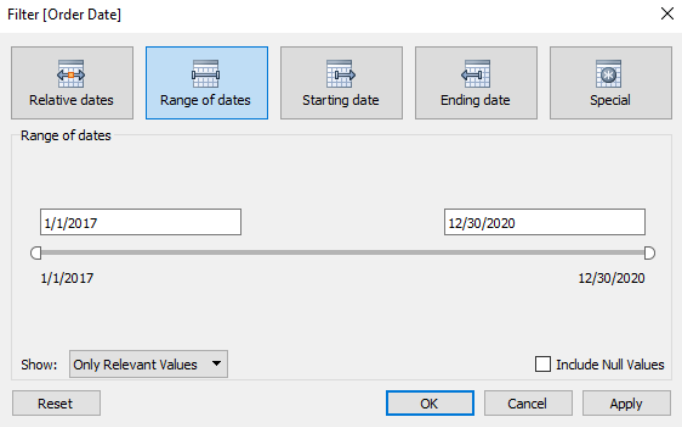
A sparkline is a very small line chart, usually drawn without axes or coordinates. It presents the general variation in values trended over time, in a simple and highly condensed way. Sparklines are small enough to be embedded in text. While the typical chart is designed to show as much data as possible, and is set off from the flow of text, sparklines are intended to be succinct, memorable, and located where they are discussed.

Here's a simple example:

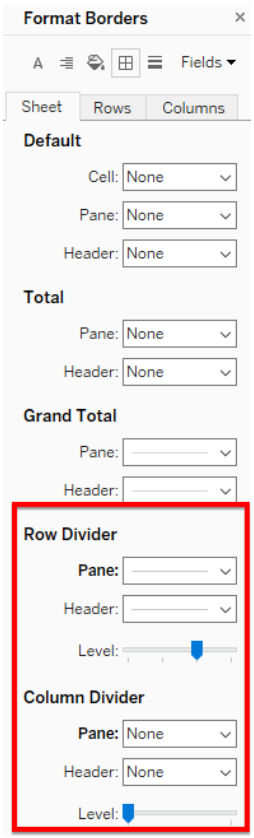


Let's build that.

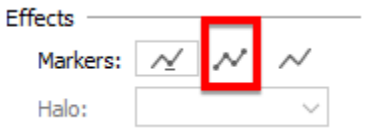
1. Drag Order Date to Columns. Set it to be Exact Date from the drop-down arrow.
2. Place Product Category on the Rows shelf.
3. Place Order Quantity after Product Category on Rows.
4. Hold down the ctrl key, click on Order Date from Columns, and drag a second instance over to the Filter shelf.
5. Click the drop-down arrow and set it to be a range of dates. Set the range to be 1/1/2017 through 12/30/2020. Click Ok.



6. Use the Size shelf on the Marks card to thin your line down.
7. Remove the axes by right clicking and unchecking "Show Header."
8. Right click on the chart itself and click Format. Then click the borders icon above the Rows tab in date pane at the left.
9. Remove the column dividers by setting them to none.
10. Hold down the Ctrl key, click Order Quantity on Rows and drag a second instance of that to the right on Rows.
11. Click the Label shelf on the Marks card and select the Line Ends option and checking the "Show Mark Labels" option at the top.
12. Title the chart "Last 90 Day Volume."



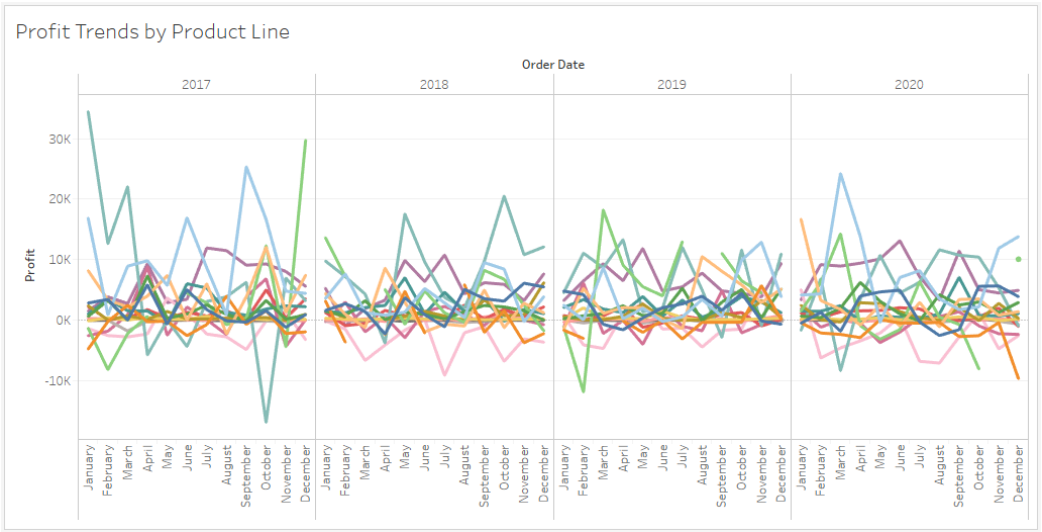
13. To get the circles to appear, click the Color icon on the Marks card, and select the middle icon.
14. You can also change the color of the line if you wish.



Note: To create the different colored dots requires making it a dual axis chart with synchronized axes, and a calculation to get the color that you want.

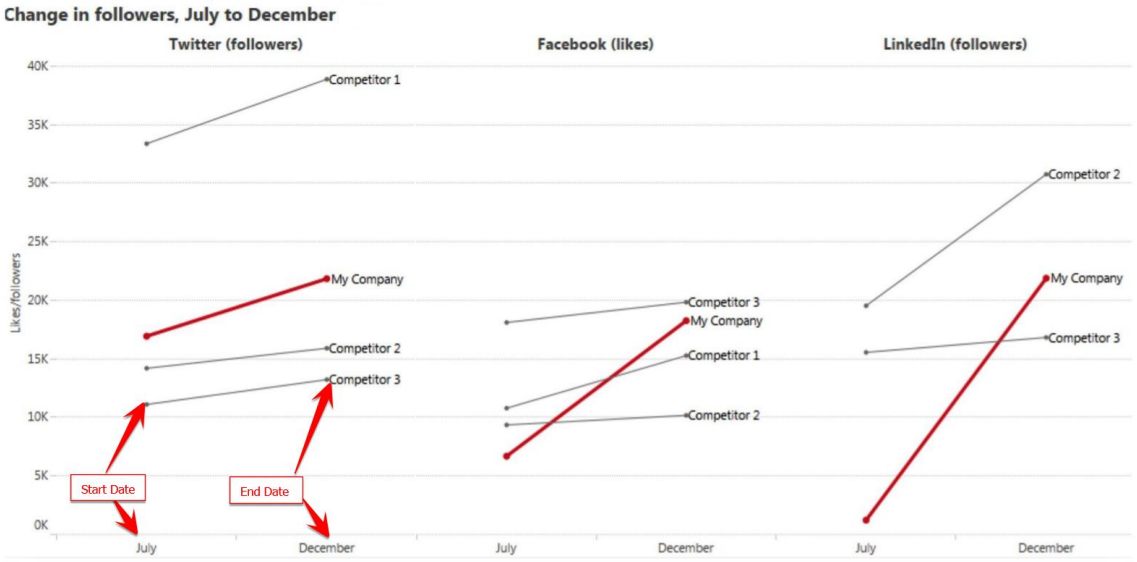
Slope Charts

Imagine we have the following chart:

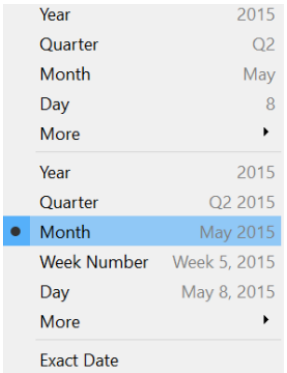


How useful is this? How much insight can we pull from this noise? Very little. With the data highlighter, that would help, but still we would have to go one at a time.

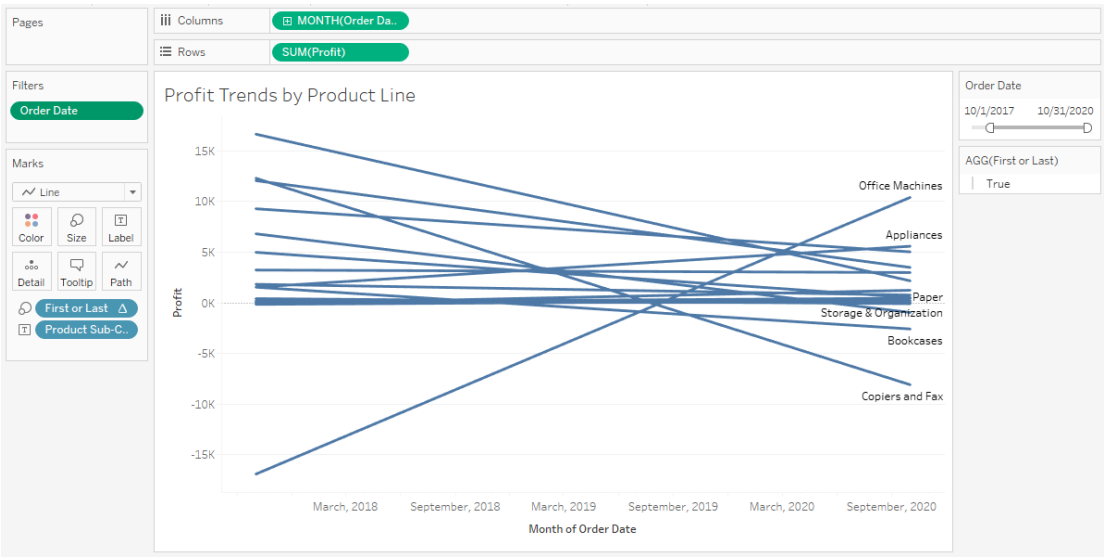
What if we just wanted to compare a start point vs. an end result and eliminate the noise in between? For this, we can use a Slope Chart. Slope Charts look like this. They show the starting point and the ending point, and they eliminate all the points in between. They're great for showing a before/after scenario, and can also reveal rankings, changes, and the magnitude of said change.



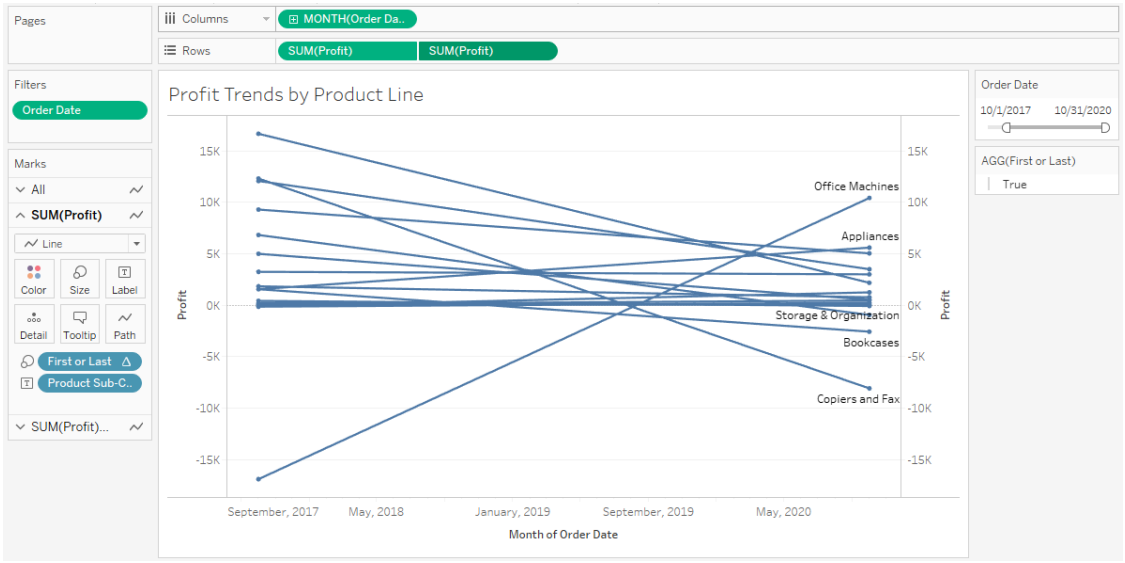
Building a Slope Chart with a Continuous Axis



1. Start by bringing Profit to Rows and Order Date to Columns.
2. Click the drop-down arrow on Order Date to Month (May 2015). It's in the second set of dates. We're making this date continuous, not discrete. The pill should now be green.
3. Bring Order Date to the Filter shelf, and set it Range of Dates.
4. Click the down arrow on this filter and choose Show Filter.
5. Now, create a calculation called "First or Last." Do this by right clicking in Measures > Create Calculated Field, or by choosing the drop down arrow from the top of the data pane near Dimensions.
6. Write the following formula: FIRST() = 0 or LAST() = 0. Choose Ok.
7. Drop this on Size. You'll see that the ends of the line get larger, while there isn't much change in the middle.
8. A size legend will appear. Right click on false and then choose Hide.
9. Drop Product Sub-Category onto the Labels shelf. Your view should now look like this:

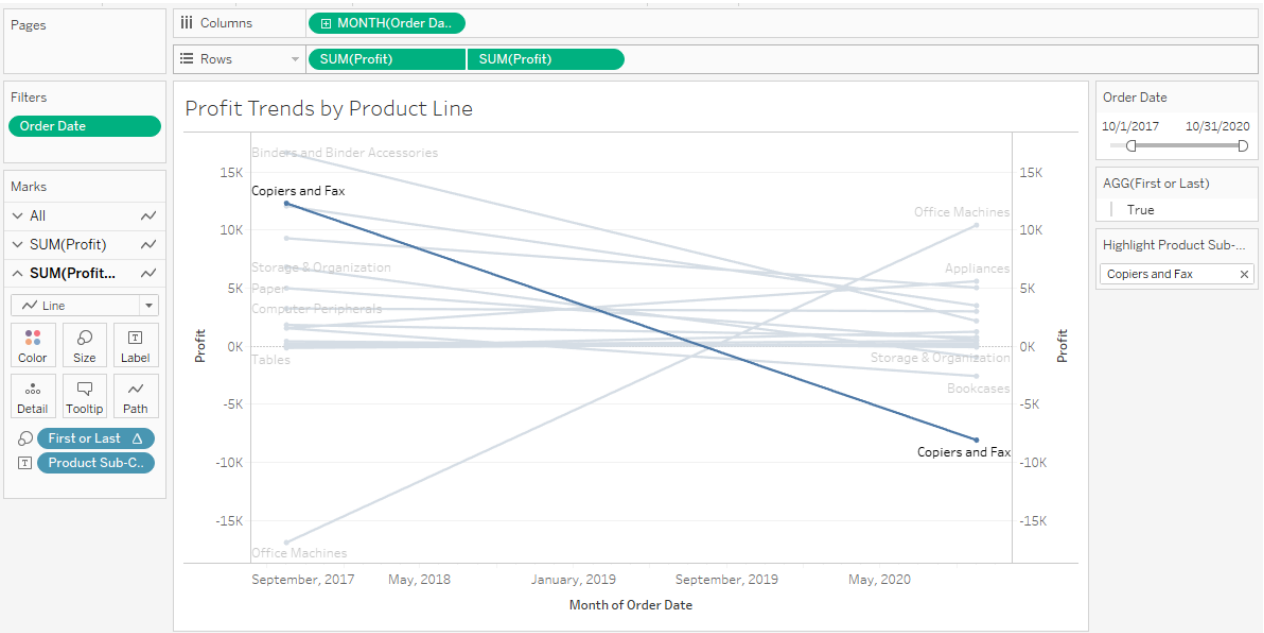


10. Click the Color shelf icon and under Effects > Markers, set it to the second icon with the dots. This will create a circle for each end of the line.
11. Hold down the Ctrl key and drag a second instance of Profit to the right on the Rows shelf.
12. Click the down arrow on that second Profit pill and choose Dual Axis.
13. Right click on that second axis and then choose Synchronize Axis.
14. Set the date field by typing in 10/1/2017 and 10/31/2020 in the Order Date filter at the top right. Your view should now look like this:



15. Click the down arrow on the Product Sub-Category pill and turn on the Data Highlighter. You can now highlight a particular element and see changes.
16. On the first, Sum(Profit) Shelf on the left, click the Label icon, and turn off "Show Mark Labels."
17. On the second one, click Label and ensure Show Mark Labels is checked. Down below that, set Marks to Label to Line Ends. Beneath that, uncheck the box that says, "Allow labels to overlap other marks." We can immediately see that Office Machines have become much more profitable for us, while Copiers and Faxes have dropped dramatically and are now losing money.

    With the data highlighter in use, we can zero in on any product line, like this:



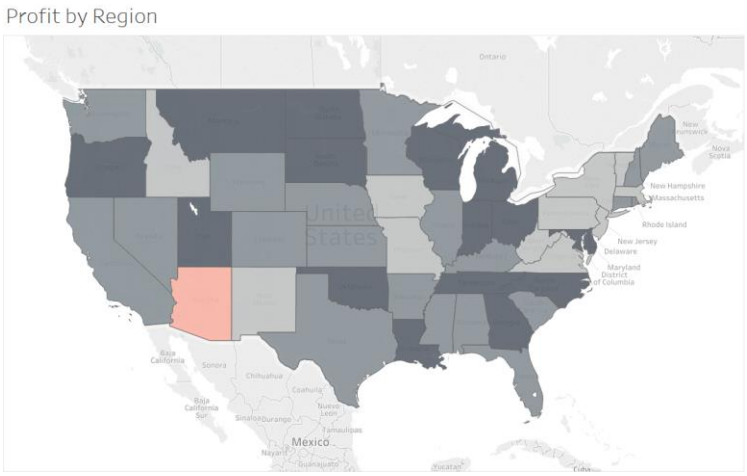
Pie Charts on Maps

We've already reviewed that pie charts should be used sparingly. But sometimes it's interesting to see a geographic representation of totals with breakdowns of those totals for smaller geo segments.

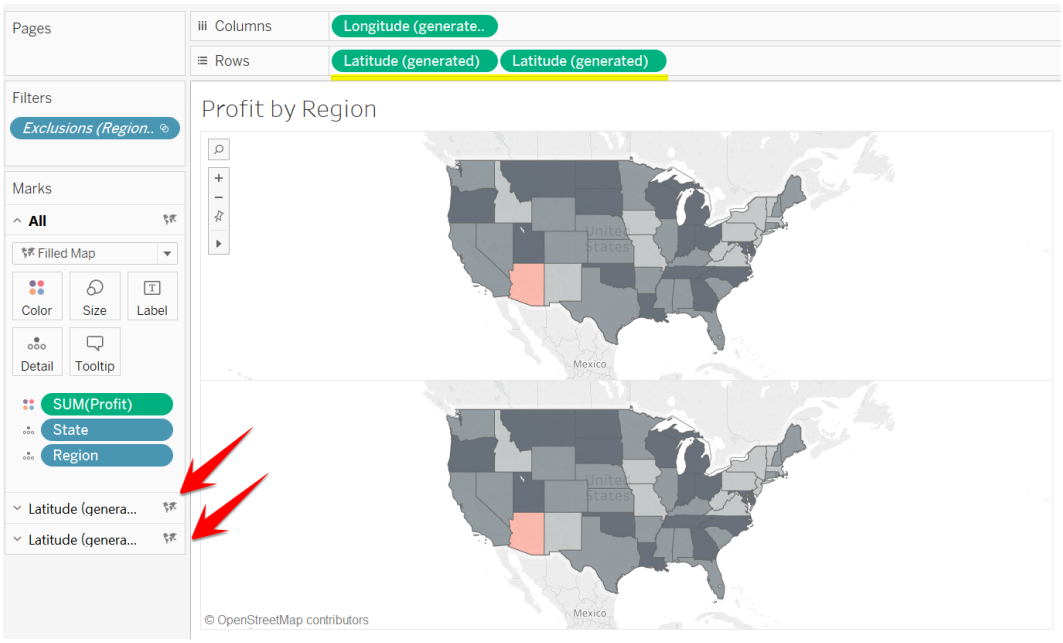
Let's say we want to see % of total profit by region on a map, and then we want to see which product categories are driving those profits in each state within a region. To do that, we could place pie charts on top of a map.

1. Bring Longitude to Columns, Latitude to Rows.
2. Then drag Buyer State out on top of the map. This zooms the map into just the United States.

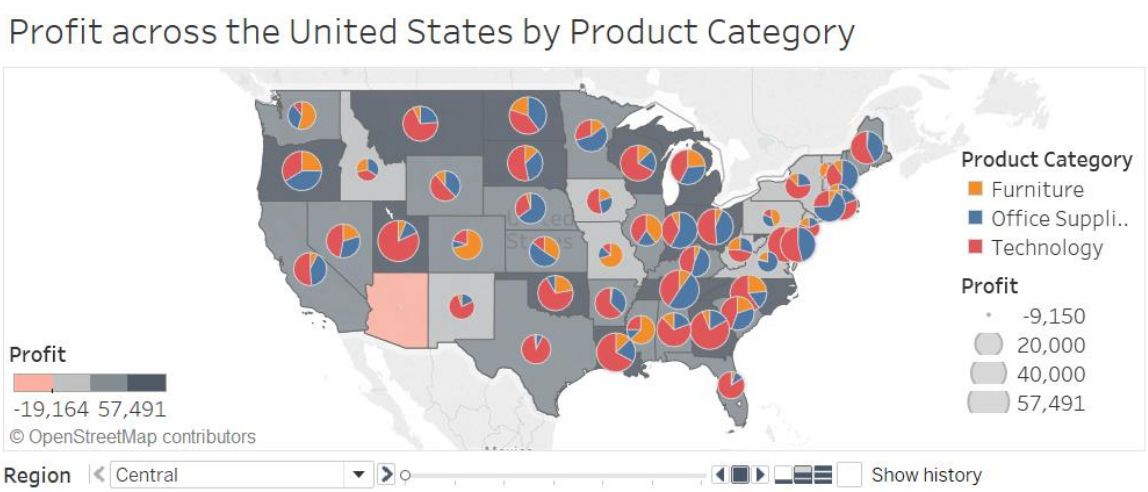
3. Drag Profit to the Color shelf.
4. Using the drop-down menu on the Marks card, change the selection from Automatic to Filled Map.
5. Edit the colors by clicking on the Color shelf and changing the color scheme to Red-Black-White Diverging.
6. Select Stepped Color and set it to 4. Choose ok.
7. Right click on and exclude Alaska.
8. Title the chart Profit by Region. You should now have a view that looks like this:

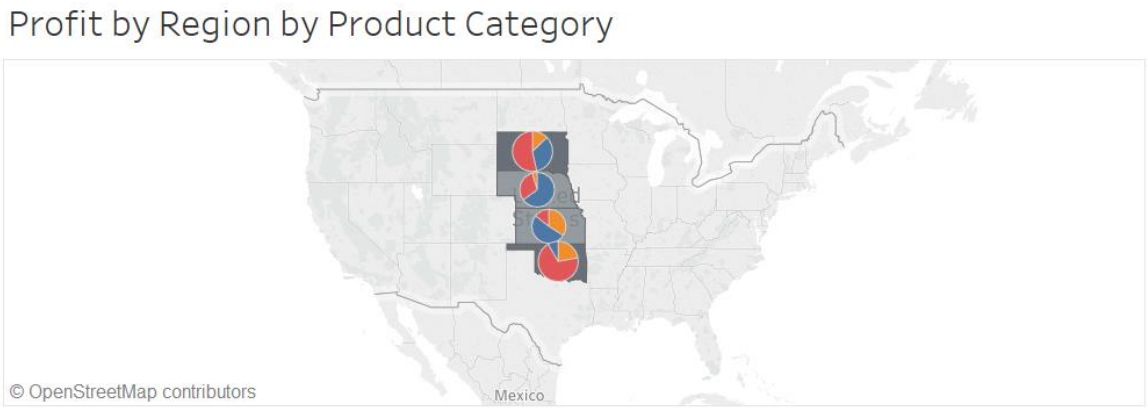


9. Now hold down the Ctrl key and click on the Latitude pill on the Rows shelf and drag a second instance of Latitude to the right. This will create another axis that we can manipulate.



10. Click the drop-down arrow on the second Latitude pill and select Dual Axis.
11. Now, click on the second Latitude shelf on the Marks card. Be sure you select that second one and that you aren't on All.
12. Set the drop-down selection from Filled Map to Pie.
13. Then, drag Profit to both Size and Angle.
14. Now, drag Product Category to Color.
15. We can now see that across the US, profits are mostly driven by Technology, but that the Midwest and North Central states are the most profitable.
16. If you wanted to take this a step further, we could right click on our sheet and Duplicate it.
17. Drag Region to the Pages shelf.
18. Build a new dashboard with both of these worksheets. Ensure the one with Region on Pages is the bottom chart.
19. Rename the top chart to "Profit Across the United States by Product Category."
20. Rename the bottom chart to "Profit by Region by Product Category."
21. Move the Pages shelf control to the middle in between the two charts and size it down so it all fits on one line.
22. Reformat the legends so that they are better spaced within the viz, and we eliminate the box on the right.





23. Click the drop down menu on Region and mouse through the different regions to determine the major profit drivers per state.

Control Charts

The control chart is a graph used to study how a process changes over time. All processes have some variability. That's normal. But large shifts or swings are cause for study and indicate something has changed about the way your process is behaving. They are used to pinpoint sources of variation.

Data are plotted in time order. A control chart always has a central line for the average, an upper line for the upper control limit and a lower line for the lower control limit. These lines are determined from historical data and typically are based on standard deviations from the average or median line in the center.

If the process is in control (and the process statistic is normally distributed, which is likely), 99.73% of all the points will fall between the control limits (usually 3 standard

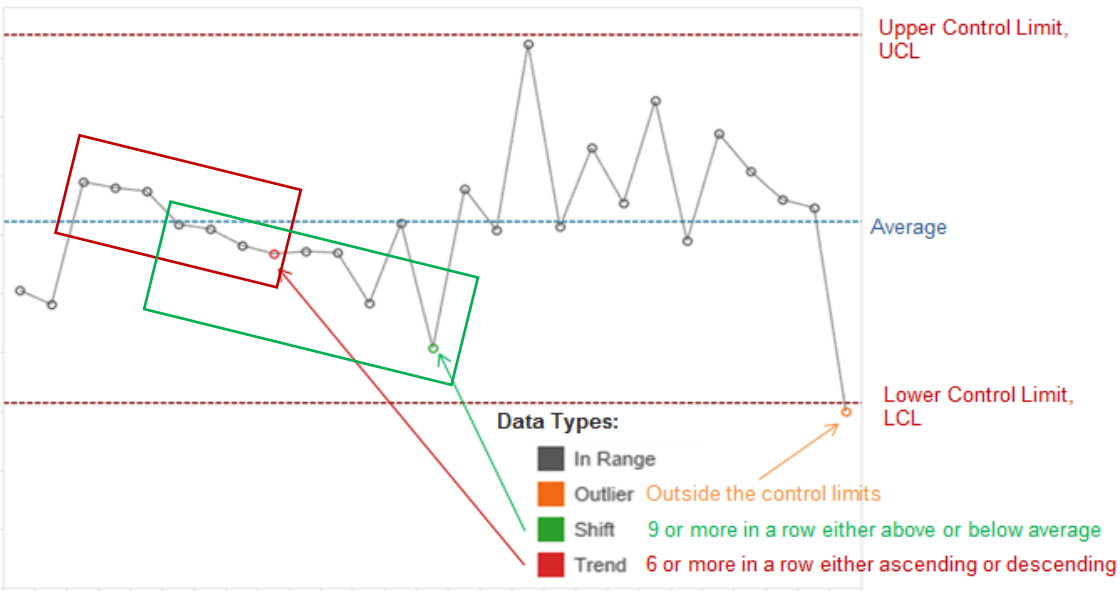
deviations above and below the average). Any observations outside the limits, or systematic patterns within, suggest the introduction of a new (and likely unanticipated) source of variation, known as a special-cause variation.

You are looking for any point above the control limit lines, or a run of 7 points that all fall either above or below the average/median line, or an upward or downward trend of 7 values. For more info, see: https://public.tableau.com/s/blog/2013/11/how-make-control-charts-tableau

To create:

1. Create a line chart.
2. Add a reference line, by right clicking on the axis or on the Analytics pane at the left, and set it to be a Line, Per Pane, and set the value to Average (or Median if desired).
3. Right click on the axis again and add another reference line.
4. Make this one a Distribution. Set it to Standard Deviation. Type in -3,3. If that seems too wide, you can also choose either -1,1 or -2,2.
5. Click ok and look for any outliers in your data.

A slightly more advanced version with color will look like this:



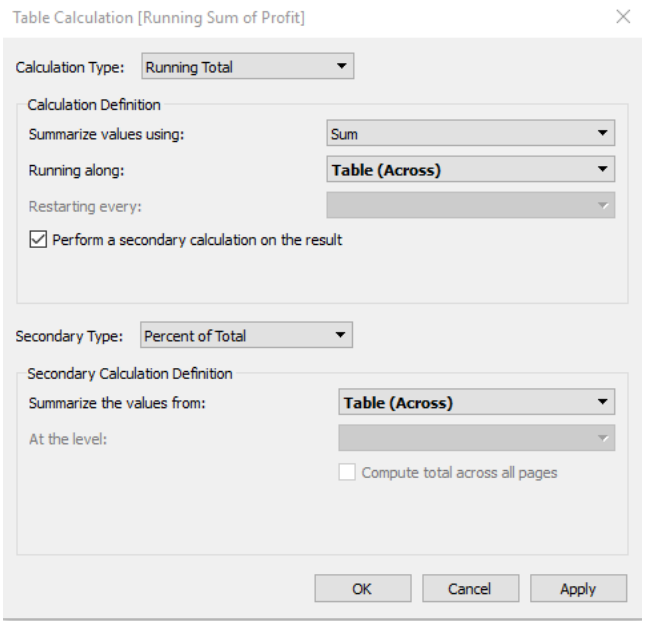
Pareto Charts

A Pareto chart is a type of chart that contains both bars and a line graph, where individual values are represented in descending order by bars, and the cumulative total is

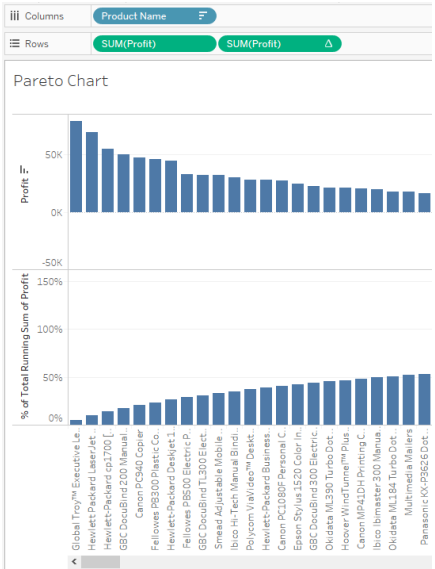
represented by the line. It highlights the 80/20 rule, which states that typically 80% of the results are caused by 20% of the factors. For example, 80% of sales are caused by 20% of salespeople. Or 80% of the problems are caused by 20% of the processes. It's helpful for knowing what's driving your issue and by how much. It also dictates where you should focus your efforts and how much impact those will create.

To create:

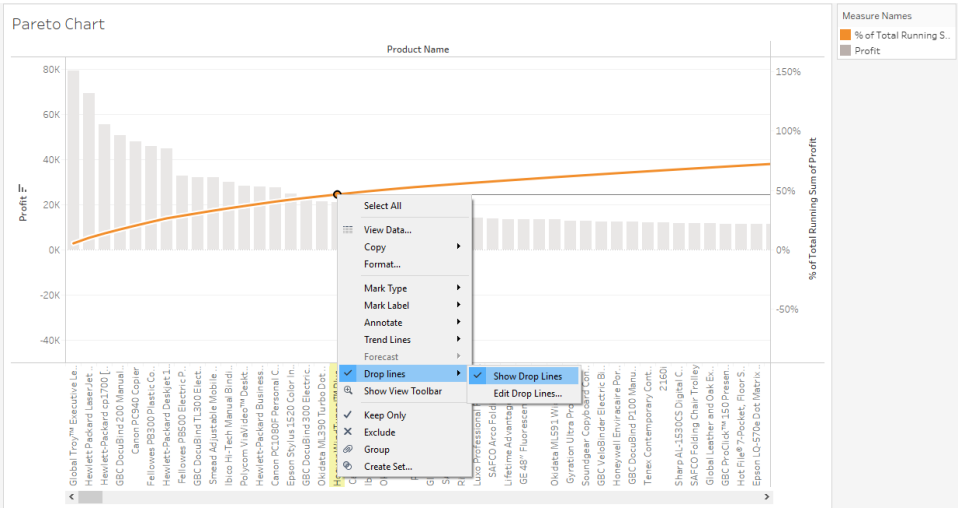
1. Drag Profit to Rows and Product Name to Columns.
2. Click the Sort Desc icon at the top below the Analysis menu.
3. Click on Profit and then hold down the ctrl key and drag a second instance of the pill to the right on the Rows shelf.
4. Click the down arrow on the of the second Profit pill. Click Add Table Calculation.
5. In the calculation type, select Running Total. Leave all the other defaults.



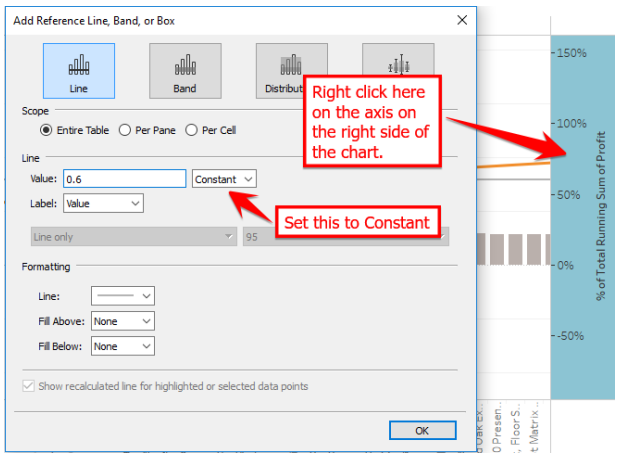
6. Click the down arrow again on that second Profit pill. Click "Edit Table Calculation."
7. In the window that pops up, check the box in the bottom left that reads "Add Secondary Calculation."
8. In the Secondary Type box, select Percent of Total. Leave the other defaults.
9. Close the box. You'll now have two charts, one on top of the other. The first has your profit sorted by product in descending order. The second shows the percent of the total profit that each individual product contributes. So if total profit was $500 and Product A had a profit of $100, it would be 20% of the total profit



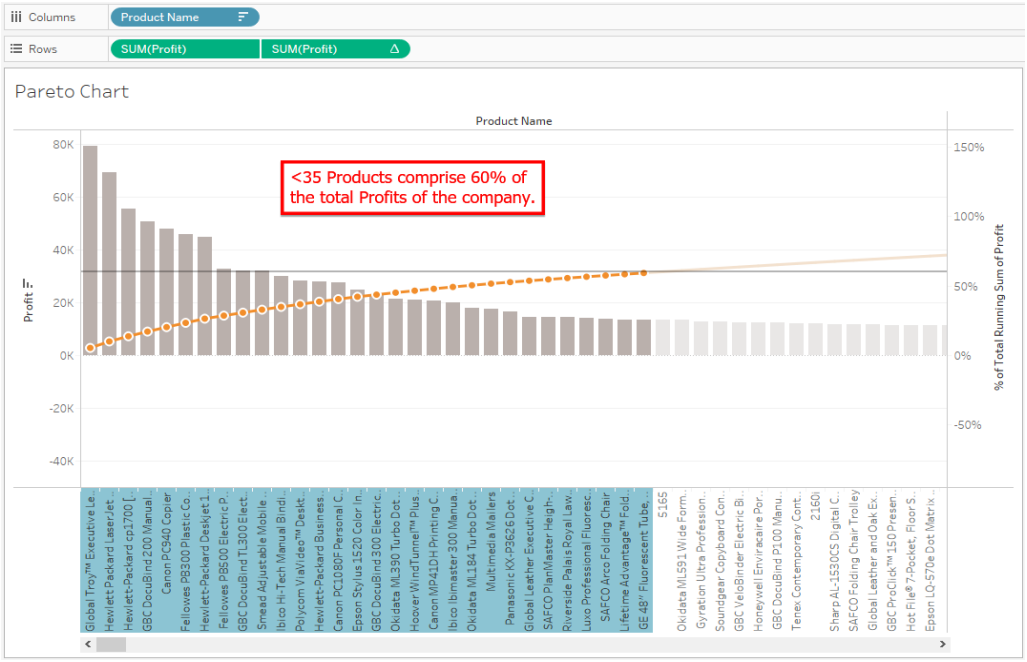
10. On the second Profit pill, right click or click the down arrow. Choose Dual Axis. This combines them into the same chart.
11. On the marks card, there are now two shelves, one for Sum(Profit), and the other says Sum(Profit) and has a triangle next to it (representing that it's a table calculation).
12. Click the first shelf for Profit and change the mark type to bar.
13. On the second shelf (with the triangle), change that mark type to line.
14. On the colors shelf, change the color of the bars to gray, and the line to orange.
15. Right click on the orange line and select Drop Lines > Show Drop Lines.



16. Right click on the secondary axis (on the right hand side) and Add a Reference Line.
17. Set the Scope to Entire Table. Set the Value in the middle section to Constant (from the drop down menu) and type in 0.6.



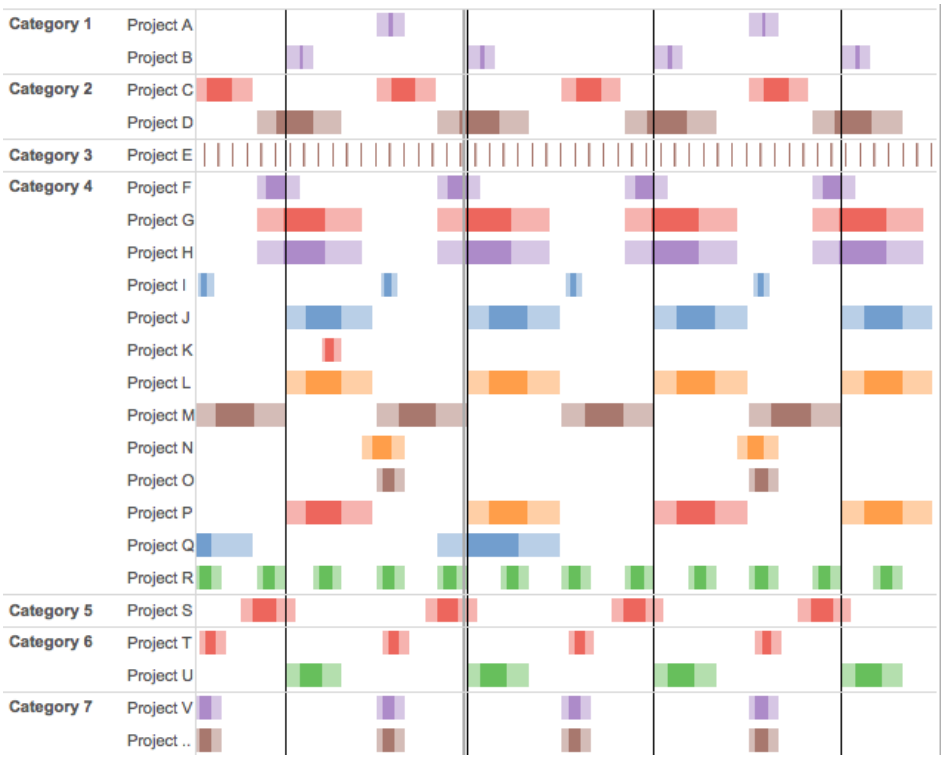
18. Click ok.
19. You can now click any point on the orange line and see what portion of the cumulative profits it represents. By lassoing the bars under the 60% reference line, we can get a count to see how many products comprise 60% of the overall company profits. Of the 1263 products in the database, <35 make up 60% of the overall profits. That's a staggering finding!



Gantt Charts and Dual Axis Gantt Charts

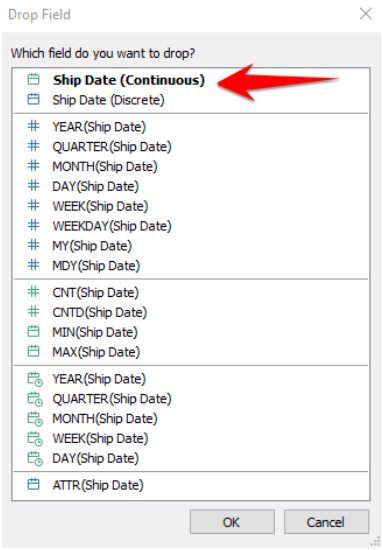
Gantt charts are very useful for tracking time in a process. This could be time from order to time to ship (important if you are Amazon Prime!), or time from hospital admit to discharge by patient or department, or any number of other examples. They are also useful for tracking progress on a project.

A dual axis Gantt chart is an even more advanced visualization. It allows you to show the peak times during a timeframe, or you can emphasize when work is being done (say, for instance, you can see when someone is logged on to their computer, and when they are on Facebook and for how long – scary huh?!). Here's an example from one of Tableau's elite Zen Masters, Matt Francis:

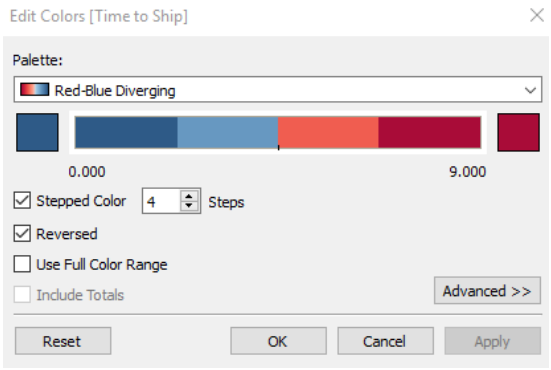


For now, we'll just build a normal Gantt chart.

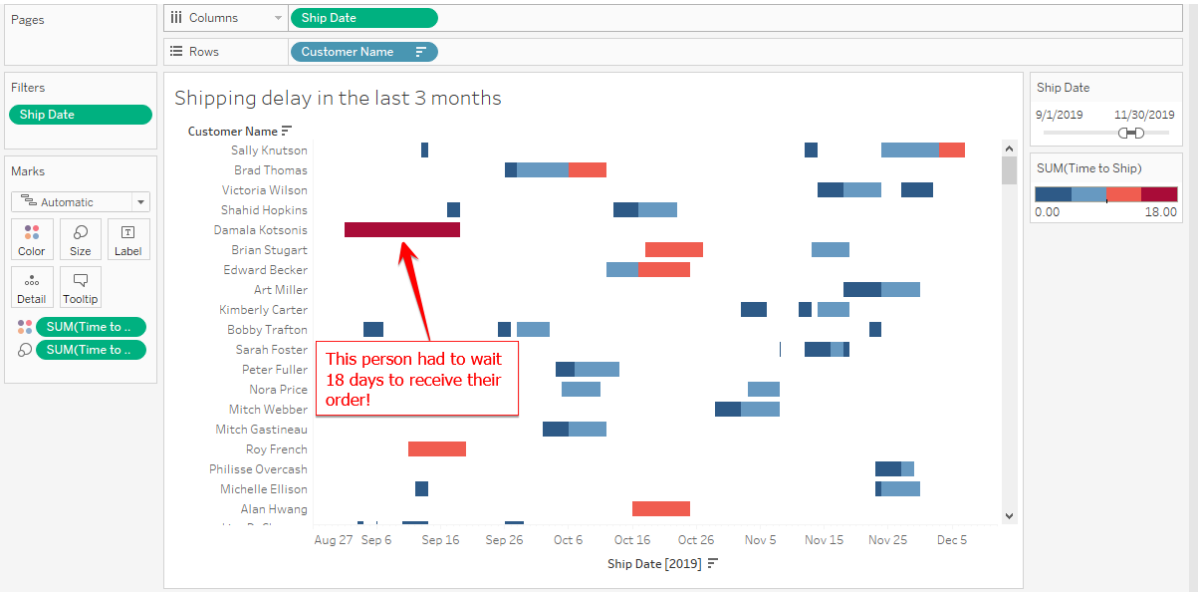
1. Right click on Ship Date and drag it to the Columns shelf. A window will pop up asking you which date type you want. For gantt charts to work, the axis needs to be continuous. Choose Ship Date (Continuous).
2.



3. Drag Customer Name to the Rows shelf.
4. The time period is too large for any type of meaningful analysis, so drag Ship Date to the Filters shelf.
5. Choose range of dates and set it to be 9/1/2019 to 11/30/2019.
6. Drag the Time to Ship calculation to the Size shelf. If you don't have this calculation, create it by right clicking and creating a calculated field with [Ship Date] – [Order Date] as the formula.
7. You now have a Gantt chart. The long bars will provide a visual indicator of who had to wait the longest for their order and will likely be an irate customer.
8. Drag this same calculation to the Color shelf as well.
9. Click on the down arrow on the color legend that appears at the top right and Edit Colors to be Red-Blue Diverging.
10. Check the Stepped Color box and set it to 4 steps.
11. Check the Reversed box. Click Ok.



12. Title the chart "Shipping Delays in the Past 3 Months."



Practice Exercises

The Quality Improvement team wants to compare Returns vs. Order Qty by Month and Year to determine where we have problems in our process. We'll build them an On-Demand chart.

1. Create a new dashboard.
2. Drag a vertical layout container from the bottom left Objects pane into the white canvas area.
3. Find your "Monthly Order Qty Trends" worksheet in your list on the left side.
4. Drag this chart onto the dashboard.
5. Delete the annotation we put there previously.
6. Now, find the tab "Orders and Returns by Sales Region" and go to that chart.
7. Put Month on the Columns shelf.
8. Move Year on the Rows shelf after Manager and before Region.
9. Use Show Me and create a heatmap – second icon, top row. Ensure that Sum(Order Quantity) is on the size shelf and that the CNT(Order ID (Returns)) is on the Color shelf.
10. Drag this worksheet into the dashboard you just created. Again, hold down Shift and place it underneath the "Monthly Order Qty Trends" chart. A gray line should appear under that chart within the blue box of the layout container.
11. Create a Filter action, titled "See Returns by Month, Year and Sales Mgr."

12. Check the first box in the top pane. Check only the second box in the bottom pane.
13. Set "Run Action On" to Menu, and "Clearing the selection will" to Exclude all values.
14. Click ok.
15. Now lasso several points by clicking and dragging so that points of the line are highlighted. Then click the link in the tooltip.
16. Set the bottom chart to Fit Width.
17. Now click back in the graph up top.
18. Label and Title the dashboard "Orders vs Returns by Month." Center the title.

Sales wants to know if they are on track to hit their daily goal. They'd like a simple visual to see how close they are to their target.

19. Build a bullet graph with Sum(Sales) on Columns.
20. Create a new parameter entitled "Sales Target Parameter."
21. Set this to Data Type: Float, Allowable Values: All. Click ok.
22. Add Order Date to the filter shelf. When the dialog box appears, choose "Month / Day / Year." Select All. Click ok.
23. Now, right click on the axis and choose Add Reference Line.
24. Click the distribution icon at the top.
25. Set the following – Scope: Per Pane, Computation Value: 60%, 80% of Sum(Sales) (get to this by clicking the down arrow and then setting Average to Sum). Leave the label set to Computation.
26. Click the Fill Above and Fill Below boxes. Set your Fill color to Stoplight.
27. Find the Sales axis label down below and mouse over it until you find an arrow. Drag this down a bit to make our chart space wider.
28. Now right click on the axis again and add another reference line.
29. This time click the line icon. Set the value to Sales Target Parameter and change the Label to Custom. Type in "Target: <Value>" (from the side arrow next to it).
30. Change the line formatting to a dashed line and set the color to dark green. Click ok.
31. Right click on the parameter at the bottom left and choose Show Parameter Control. Now type in some values of the box and see how the bullet graph rescales.
32. Click on the down arrow of the date filter and choose Show Quick Filter. Adjust the date and see what happens to the chart. You will see more noticeable

differences if you grab a range of dates by selecting one date, holding Shift, and then clicking on another date.

33. Click on the line reference line and choose Format. Set the label alignment to be Top (via the three stacked line icon).
34. Click on the 60% reference line and format the label alignment there to be left.
35. Adjust the width of the bar by using the Size shelf to make it smaller.
36. Click Color and make the bar a gray color.
37. Rename the tab "Sales Bulletgraph."

If time permits, build another view showing which products generate most of the sales. This will be a Pareto chart.

38. Create a new tab. Name it "Sales Pareto."
39. Bring Sales to Columns and Product Name to Rows.
40. Sort descending by Sales.
41. Hold down the Ctrl key and click on Sales and drag a second instance of Sales to the right of the first Sales pill.
42. Click the down arrow and add a quick table calculation for Running Total.
43. Right click the down arrow again and Edit the Table Calculation.
44. Check "Add Secondary Calculation" and make it a Percent of Total, compute using Table (Down). Click the X at the top of the window.
45. Right click on the second Sales pill again and make it a dual axis now.
46. On the Marks card, make the first Sales pill a bar from the drop down menu. Make the second Sales (with the triangle icon) a line.
47. Color the bars gray and the line orange.
48. Right click on the orange line and add drop lines.
49. Swap Rows and Columns so that your chart reads left to right not top to bottom.
50. Count the first 10 bars and determine what percentage of the total sales they comprise.

Pop Quiz (answers in Appendix)

1. If I wanted to compare profit to sales within the same chart, what would be the best chart to use?
    a. Pie chart
    b. Bar in bar chart
    c. Control chart
    d. Gantt chart

2. If I wanted to show length of time for a series of tasks within a project, or show amount of time between two events, what would be the best chart type to use?
    a. Pareto chart
    b. Sparkline
    c. Bar in bar chart
    d. Gantt chart
3. If I wanted to see an end result vs. a starting point and eliminate distracting data, would be the best chart to use?
    a. Sparkline
    b. Control chart
    c. Slope chart
    d. Bullet graph
4. If I wanted to see how we compare against a target value, or how close we are to a certain goal, which chart should I use?
    a. Pareto chart
    b. Gantt chart
    c. Bullet graph
    d. Gauge that looks like a thermometer I brought in with clipart

---

Clustering – What is it?

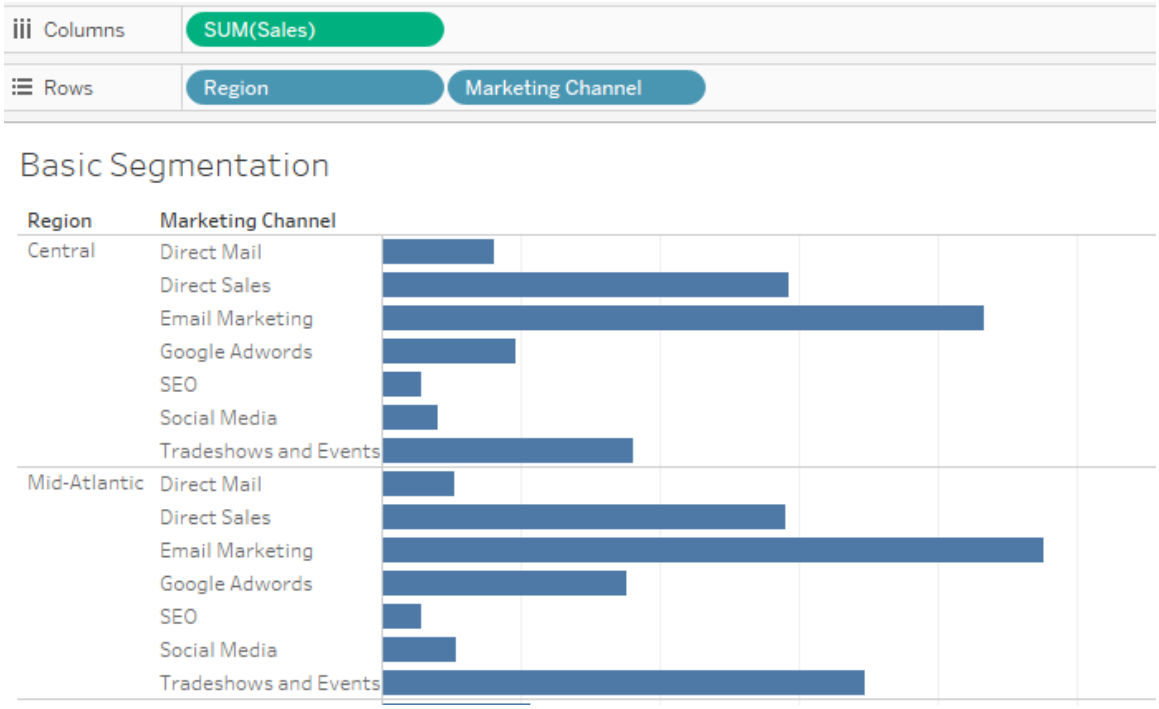
Clustering is a new feature in Tableau 10, and puts advanced statistics into your hands with a few clicks. Clustering allows you to easily identify statistically similar groups. In plain English, you can throw a bunch of dimensions and measures together, and determine similarities and create look-a-likes, which you can then drill into for more detail or compare how each group behaves relative to each other.

Just as we've said that sorting values and identifying your top and bottom values are key skills of a good analyst, the ability to segment data into useful groups or bins is a must in your toolbox. Clustering takes that to a whole new level. You don't need code or need to be a trained statistician to access it. And, like the rest of Tableau, it follows the now familiar pattern of viz, analyze, iterate, enabling you to find insights quickly.

So, let's look at an example: customer segmentation.

1. Let's build a view with Marketing Channel on Rows and Sales on Columns. This segments all revenue via the marketing channel by which we acquired it.
2. Now let's add Region to the left of Marketing Channel on Rows. We've now got two macro-segments in the view, with the breakdown being the combo of each.

So, we can now answer questions like: “how many sales ($) did we acquire via Direct Mail in the Central region?”



But, wait. Can’t we already create segments? Why do we need Clustering?

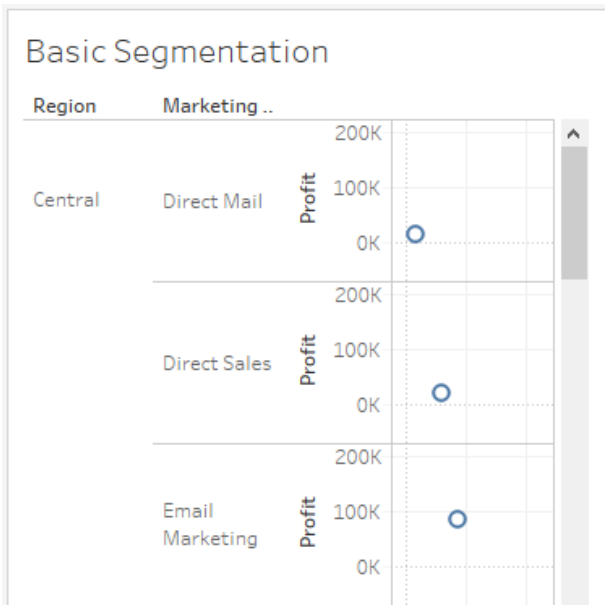
We can already segment the data and the chart however we want, right? We could continue adding Dimensions (aka segments) to the Rows shelf, if we wanted, and get ever more granular segments. All true points.

But, what if we wanted to add more variables and determine the interactions of those variables? For instance, we might wonder: “how do these 6 things interact together and what results do they produce?” Or what if we wanted to add Measures instead of Dimensions? For example, purchase patterns (Sales) and amount we actually make (Profit) and return or discount patterns (Discount, Returns).

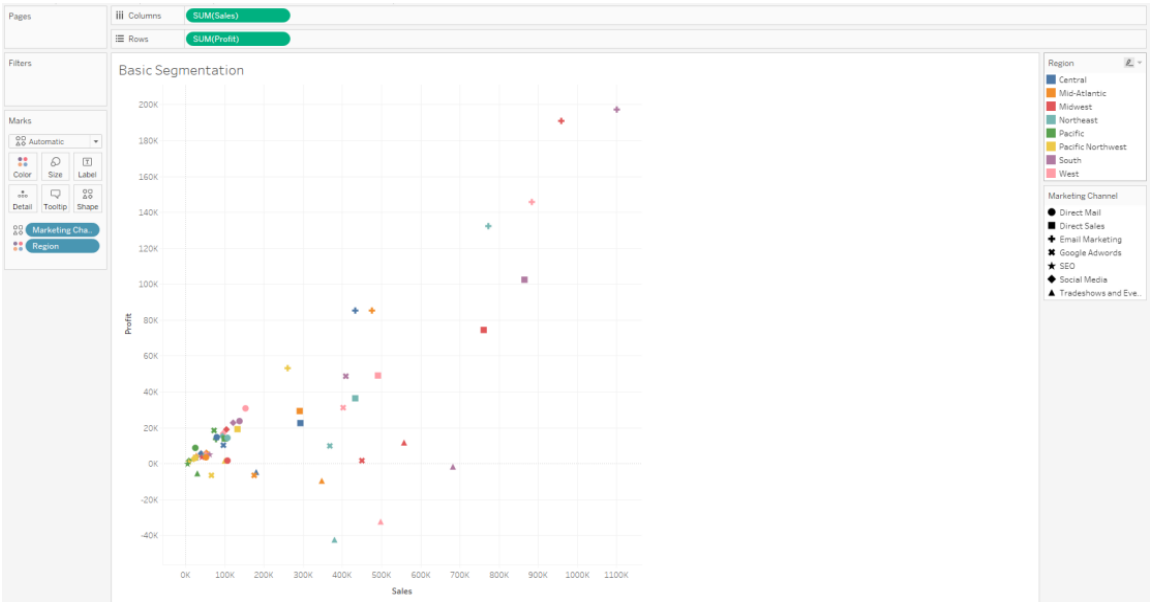
Clustering allows us to add this additional information. This helps us move beyond simple segments to advanced, incorporating data on behavior patterns and actions (Measures), as well as attribute information like Region or Marketing Channel (Dimensions).

How do we create clusters?

3. Returning to the same chart we created above, let’s now add Profit to Rows.
4. You should have several scatterplots, each with one dot in it.

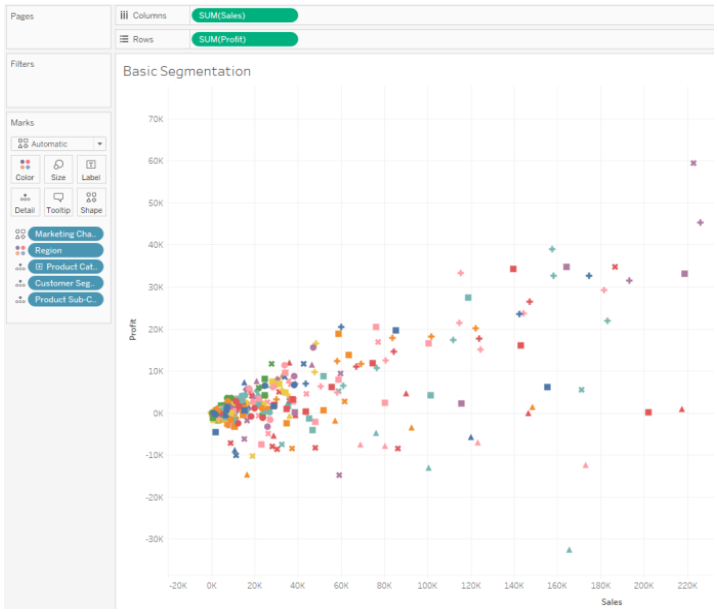


5. Click Show Me up at the top right and choose the Scatter Plot option to get this into more of a useful format. You'll see that Marketing Channel and Region are on the Shapes and Color shelves, respectively.

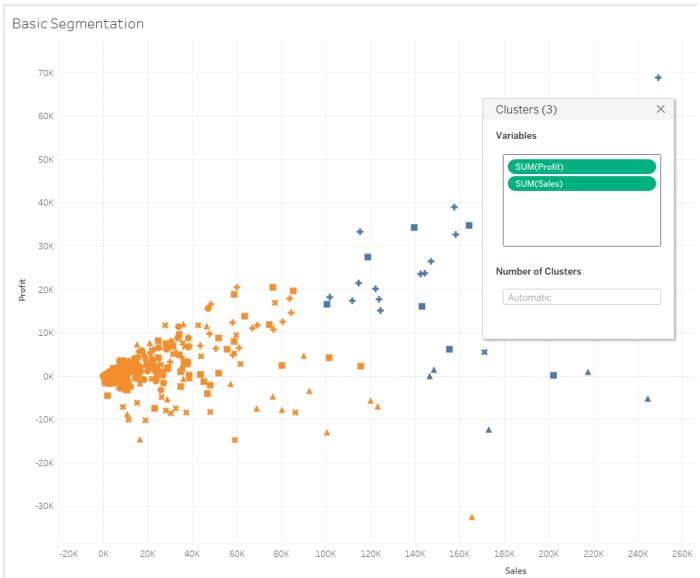


6. Set it to Entire View from the drop down menu at the top.
7. Now, let's add several more Dimensions. Add Product Category, Customer Segment, and Product Subcategory to the Detail shelf.

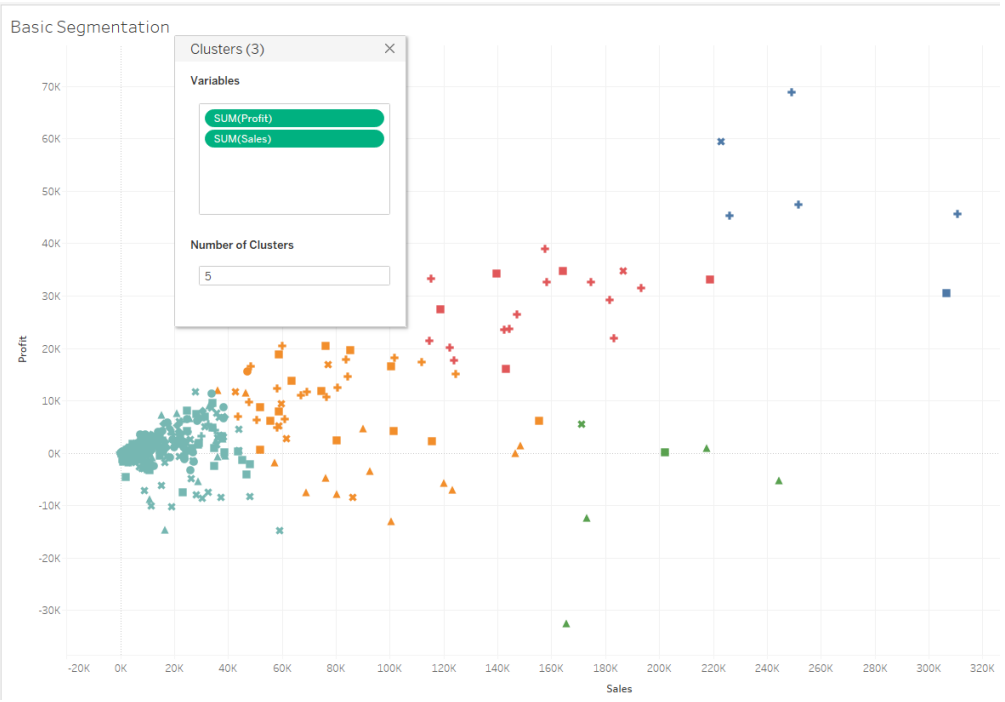
8. Click on the Shapes card to set each of the marks to Filled from the drop down menu labeled “Select Shape Palette.” Choose Assign Palette and click Ok.



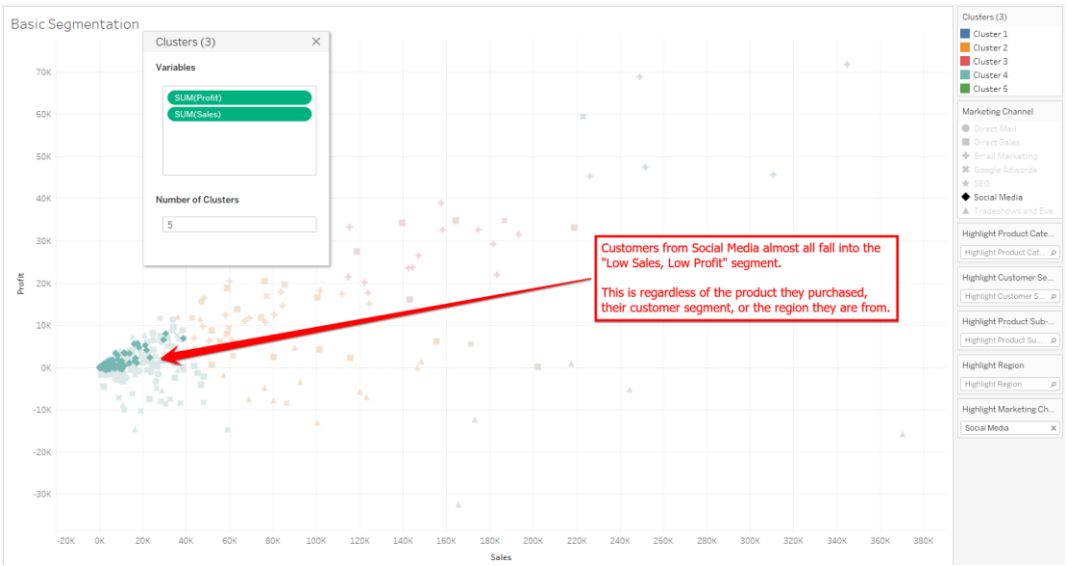
9. Now, click on the Analytics tab at the top left, above your Dimensions.
10. Click Cluster and drag it out. Be sure to place it on top of the Cluster box that appears.
11. Notice that 2 clusters are generated automatically from the data.



12. Let's play with the number of potential clusters. Change the number from Automatic to 5. You should now see the different colors.



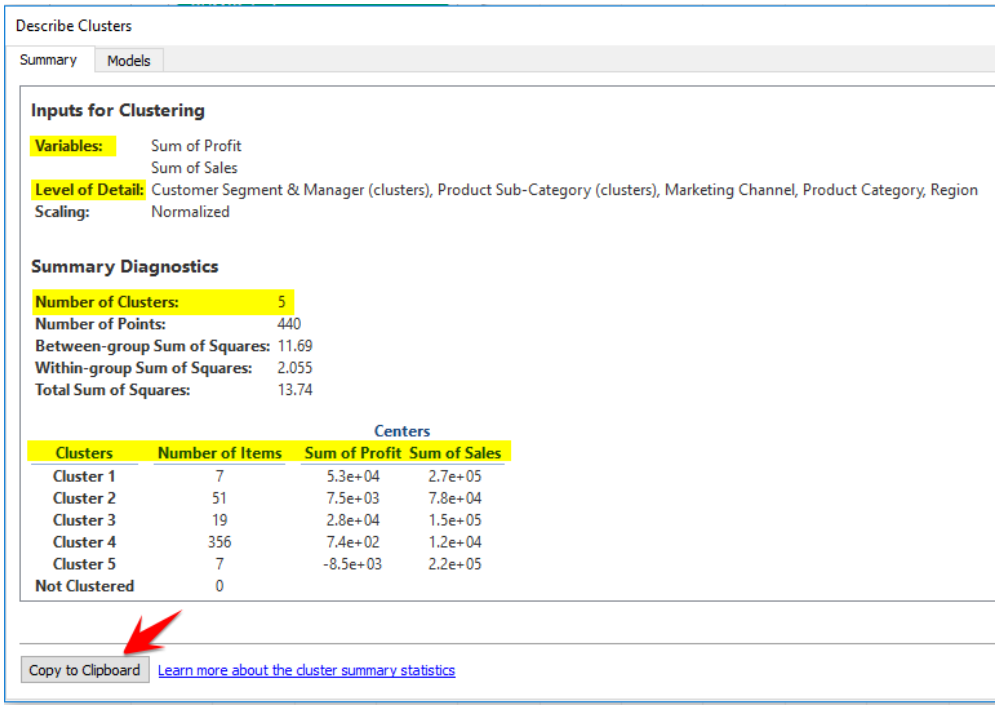
13. Go over to top right where the data highlighter shows the different clusters. Click each one in succession to highlight that segment on the scatter plot. Are you seeing some interesting groups, like a group of high sales, low profit?
14. Click on the down arrow on each pill that you put on the Detail shelf and select "Show Highlighter."
15. These should appear on the right hand side. Click through these to see if there are any interesting insights that emerge. For example, under the Market Channel highlighter, choosing "SEO" or "Social Media" reveals some interesting insights. Or choosing "Google Adwords" reveals an interesting outlier.



How do I understand each of my clusters beyond just eyeballing it?

16. Click the down arrow on the Clusters pill which should be on your Color shelf.
17. Choose Describe Clusters.
18. A window will appear with a lot of information about how this was created. You want to pay attention to the following:
    a. Variables – these are the measures that you are crunching to find look-a-likes (i.e. group similar customers by sales and profit)
    b. Level of Detail – these are the dimensions that you're incorporating into the cluster (i.e. show me look-a-like customers by sales and profit, by analyzing customer segment, marketing channel, product category, etc. and finding commonalities across all of those).
    c. Number of clusters – these are the distinct groups or segments that the algorithm found
    d. Clusters – you need to scroll down to find these.
        i. Number of Items – shows how many data points are in each cluster (these could be your bars or the circles on a scatter plot)
        ii. Centers – this is the average value within each cluster. You'll see the obvious differences.
            1. Note: it's ok to have clusters of different sizes as data may group more strongly at one end then another, but you want each cluster to have enough data points to be meaningful.
            2. If it only has one or two, consider excluding those from the view as they might be outliers skewing your results, or consider changing the number of clusters.

3. Note: Most of the cluster centers will appear in scientific notation, which is frustrating. If you click the Copy to Clipboard button and paste it into Excel, you can format the numbers so you now correctly what they represent.



How do I explain this to other people?

In English: Find members of a potential group (could be customers, could be cities, could be anything you're trying to group on) that are as similar to each other as possible, and as dis-similar as possible to the next group. We want each group to be as unique and distinct as possible, while we want each member of a particular group to be as similar as possible.

Quantitatively: For a given number of clusters or look-a-like groups (denoted by the letter "K"), the algorithm partitions the data into that many clusters or groups. The algorithm will determine what it thinks is the optimal number of clusters for you, based on your data. But you can easily change that to see if new patterns emerge. Each Cluster has a center (centroid) that is the average value of all the points in that cluster. Each cluster is a valid statistical grouping that will update dynamically as data values change or as new data is added.

An example: Let's say you have information about four Domino's pizza chains, and a list of customer addresses. But those customer addresses aren't tied to any particular

Domino's location. You'd have to manually sort through the addresses and compare them on Google Maps to determine which location they should order from. Clustering does this automatically. It would crunch through the data and then determine which neighborhoods are around each Domino's location. You'd have four clusters. This is essentially what Google does when you search for "pizza near me," by the way.

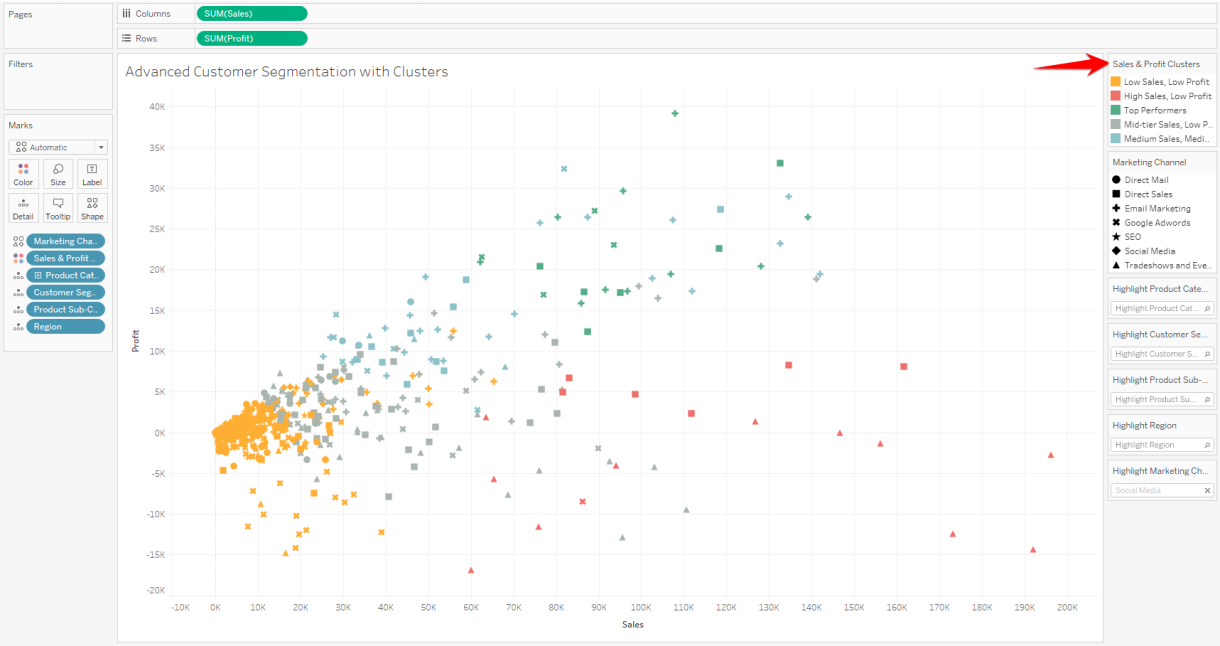
How Do I Reuse These Clusters Again, or in Other Analyses?

You'll notice that the clusters don't have very descriptive names, i.e. Cluster 1, etc. We're going to change the names of these by turning these Clusters into groups in our Dimensions pane that we can reuse.

19. Choose Close from the Describe Clusters window.
20. Hold down the Ctrl key and then click on the Clusters pill on the color shelf, and then drag this over into Dimensions.
21. Now, double click the Clusters pill you just dragged into Dimensions and rename it to "Sales & Profit Clusters." This is now a field that we can reuse again later, which will be very helpful in analyzing certain segments of customers.
22. Click the down arrow on the renamed pill and choose Edit Group.
23. Right click on Cluster 1 and choose Rename. Type "Low Sales, Low Profit."
24. Follow the same procedure for Cluster 2 (note that they may not be in numeric order!). Rename is to "High Sales, Low Profit."
25. Rename Cluster 3 to "Top Performers."
26. Rename Cluster 4 to "Mid-tier Sales, Low Profit."
27. Rename Cluster 5 to "Medium Sales, Medium Profit."
28. Now drag the updated "Sales and Profit Clusters" pill and replace the existing Clusters field on the color shelf. You can do this by placing this pill directly on top of the other one. Or, by dragging the current field on Color off and replacing it with the new.
29. Now, let's change the color scheme, so that our colors convey a little more meaning. On the legend, click the drop down arrow at the top right, and choose Edit Colors.
30. Set the color palette to Superfishel Stone in the drop down menu.
31. Now choose the "Top Performers" segment and the click on the dark green pill.
32. Repeat this procedure and change "Low Sales, Low Profit" to the orange color. Change "High Sales, Low Profit" to red. Change "Mid-tier Sales…" to the light olive color. Change "Medium Sales" to the aqua color.
33. Choose ok.

a. We now have some statistically valid segments that we can reuse and that are highlighted with meaningful titles that indicate the next step. For example, "High Sales, Low Profits" leads us to the very obvious "why" question. We can then drill down deeper to see what else surfaces from these data points that indicate actions we need to take.

34. Rename your sheet "Advanced Customer Segmentation with Clusters."



What are some common use cases where I could use clustering?

Customer Segmentation: Say you have a group of customers that logs in very infrequently, never calls support, started with low monthly recurring revenue, but spent tons on upgrades over time. That's an odd group with tremendous organic growth and low costs, even though initial revenues were low. Clustering can find groups like this.

Market research: How do we determine different groups in the market and create products and marketing messages that resonate with those people? For example, a bank found a group of entrepreneurs that was using equity from their homes via a 2nd mortgage to fund their startups. Knowing that led to a whole new line of products for the bank that resonated much stronger with that group.

Customer Surveys: What clusters crop up among satisfied customers, what clusters crop up among unsatisfied customers? Are the unsatisfied customers also utilizing your excellent support services?

Matching or Recommendation Algorithms, like Netflix: For example, based on movies that have a Strong Female Protagonist, Witty Humor, and British Actors, we recommend all movies based on every Jane Austen book ever.

Telecom: Position the cell towers so that all customers receive optimal signal strength based on addresses, usage patterns, roaming, subscriptions, peak times, traffic patterns and roads, etc.

Scheduling: Say you're a police chief trying to maximize your officer time with limited budget. You need to schedule patrols at peak times in the most crime-likely areas, again based on any number of factors, like time of day, weather, income and education levels, past crime events, types of crime, known gang locations, etc.

Practice Exercises

The VP of Sales is back for more. This time she wants to see how her sales managers are performing. She'd like to know what's affecting each manager's contribution to the bottom line – i.e. order quantity, discounts, pricing, customer segment and margins. You know for an analysis that complicated, you're going to need to use Clustering.

1. Create a scatter plot with Sales on Columns and Profit on Rows.
2. Set to Entire View.
3. Bring out Manager to Detail. (This should create a scatter plot with 7 marks on it).
4. Now, bring out Customer Segment to Detail.
5. Add Order Quantity to Detail.
6. Now, bring out Product Base Margin, Discount, Unit Price, and Discounted Price, all to Detail. Ensure each of those is set to Avg.
    a. Hint: Unit Price may be a Dimension, but you can still right click and drag it in and set it to Avg and it will convert it back to a Measure for this viz.
7. Create a Cluster.
    a. Hint: Use the Analytics tab above Dimensions.
8. Change the number of clusters to 6.
9. Set the Shapes to Filled (remember to choose Assign Palette in the window that pops up).
10. Size up the size of the marks just a touch.
11. On your Manager legend at the top right, ensure the Pencil icon is selected so that you can highlight the sales managers.
12. Hold down the Ctrl key and drag the Clusters pill from Colors into Dimensions.

13. Rename this pill in Dimensions to “Sales Manager Clusters.”
14. Give each cluster a more descriptive name by right clicking on the Sales Manager Clusters pill and Edit Group. Double click each cluster name and rename them to be the following:
    a. Note: Normally, you’d need to copy these to Excel and look at these to determine what’s causing the clusters.
    b. In this case, I’ve given them to you:
        i. Cluster 1 – High Price, Low Discount, High # Orders
        ii. Cluster 2 – High Price, High Discount, Highest # Orders
        iii. Cluster 3 – Mid Price, Lowest Discount, Lowest # Orders
        iv. Cluster 4 – Low Price, High Discount, Mid Range # Orders
        v. Cluster 5 – Mid Price, Highest Discount, Lowest # Orders
        vi. Cluster 6 – Highest Price, High Discount, Mid Range # Orders
15. Now, drag it back to the Color shelf. You’ll notice the title of your legend changes over on the right.
16. Let’s now create a color scheme that maybe provides a visual indicator of whether a segment is good or not. Click the drop down arrow from the legend and Edit Colors. Give the clusters the following colors from the Summer Palette.
    a. High Price, Low Discount, High # Orders - Aqua
    b. High Price, High Discount, Highest # Orders – Dark Teal
    c. Mid Price, Lowest Discount, Lowest # Orders – Dark Red
    d. Low Price, High Discount, Mid Range # Orders - Green
    e. Mid Price, Highest Discount, Lowest # Orders - Pink
    f. Highest Price, High Discount, Mid Range # Orders - Orange
17. Title the Sheet “What’s Driving Sales Managers’ Performance.”
18. Add reference lines to show Avg for both Sales and Profit.
19. Now click some of the sales manager’s names from the color legend at the top right and see who needs to improve.

Storypoints

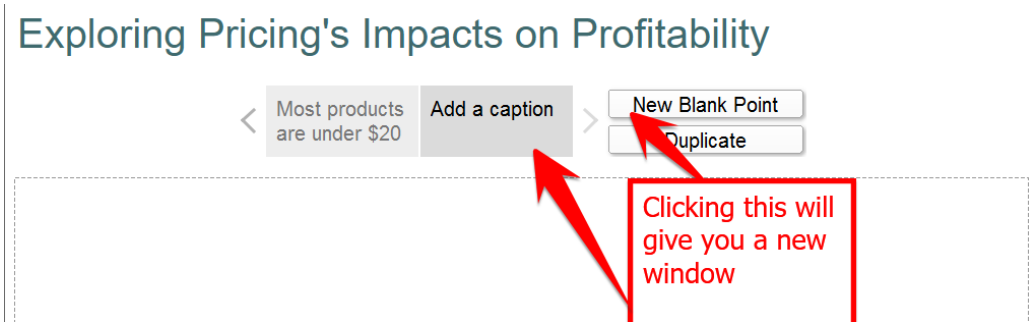
Storypoints are a powerful, often underutilized, feature of Tableau. Most people are used to seeing tables of numbers from Excel, with maybe a chart or two. It’s a manual process to create and maintain those, and they aren’t always done well. Further, an Excel chart can encode one or two values at the most. A Tableau chart can encode several – at least 6+ variables. When you place several charts like these in a dashboard, it can be overwhelming.

Storypoints, along with Actions, can help solve this problem. Just like Powerpoint has the animations feature, allowing you to have one bullet point appear at a time, you can gradually unfold your analysis step by step using Storypoints so that your end user doesn't get lost. The end goal is to have them arrive at the same conclusion you did. Storypoints takes them through the process step by step with you.

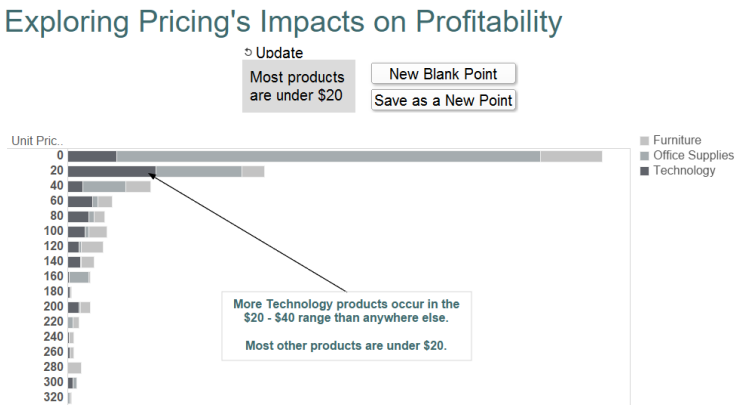
To create a dashboard using Storypoints:

1. First, ***and most importantly***, think of your analysis as a story. It's got to have a start point, a series of steps that move you forward, a climax – a key insight or conclusion you want your audience to take away – and finally a resolution or next steps.
2. If you get this step right, then using Storypoints effectively will fall into line.
3. Click on the icon at the bottom far right entitled "Create Story." This will create a new tab entitled "Story 1."

4. Add a title to the Story that will apply to your entire dashboard. This is equivalent to a book title; it covers the entire story.



5. Drag your first chart, dashboard, text box, image or webpage, onto the dashboard canvas.
6. Make sure there is a title or captions on your charts and images so that this makes sense to your end user. Think book chapter title here.
7. Double click in the "Add a Caption" box at the top and add a descriptive sentence that helps build the action. Think Tweet length here – 140 characters or less.



8. Then click enter. A new Storypoint will appear. This is equivalent to animating in your second bullet point in PowerPoint, or adding a new chapter to your book.
9. Repeat the same steps as above as you build your story.
10. To format your Storypoints or Title, click on the Story menu link at the top and click format.

Following these steps will transform a complicated analysis into more bite size chunks, which will help your end user get to the place you want them to. If done right, it can lead to big action!

Performance Tuning

For Excel files that aren't too large, you probably won't need these techniques. But as your files get larger (both data sources and your workbooks with calculations), and especially as you are getting data from your company's servers, performance can really slow down. Since people are accustomed to getting what they want from the web instantly, having slow workbooks can be detrimental, causing people to move on without consuming your content.

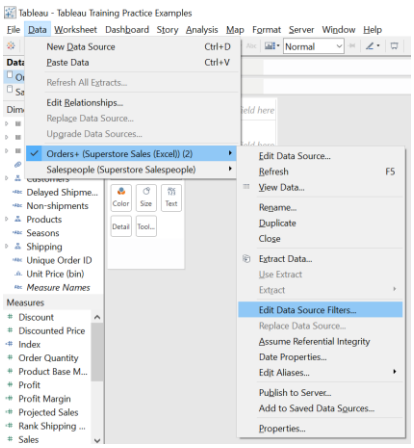
Tableau has a built in workbook performance grader that enables you to analyze your workbook's performance. To utilize this effectively,

1. Open a brand new, blank Tableau workbook.
2. Go to Help > Settings and Performance > Start Performance Recording.
3. Now open the workbook that you want to optimize. This will record how long it takes for all the individual charts and dashboard elements to load initially.
4. Now click through your workbook – filters, actions, different dashboards, Storypoints, etc. Utilize all the functionality you've built into the workbook.
5. Once done, go back to the Help menu and choose stop recording.
6. Tableau will automatically create a new workbook with the stats of your performance recording. It will show you what took a long time and where the hold ups in the process were. If your data is a Live connection, it will also show you which calculations run slowly.

7. For more information on how to interpret this, go to the following link: https://onlinehelp.tableau.com/current/server/en-us/perf_record_interpret_server.htm

Some Performance Best Practices

- Use extracts as much as possible.
    - If generally working at a higher level of aggregation than the level of grain of your data, use your extract to aggregate at a higher level to boost processing times. It's less overall data for Tableau to crunch through when you are building views.
- Use data source filters to reduce the amount of data you are bringing in. i.e. Eliminate fields you don't need.



- Minimize the use of Quick Filters. Use Actions and Parameters as much as possible.
- Try to limit the number of tabs in any given workbook to less than 40. If you'll need more than this, consider building two workbooks and linking them with URL actions.
- Push as many of your calculations as possible into a data source or data connection.
    - This could be a calculation in your SQL or Excel file, or it could be in an extract that you've published to Server, and then connect to in your workbook you are working in.
- Avoid string calculations where possible.

    - Tableau evaluates integers the fastest, then Boolean (true/false) values, then strings.
    - Often string calculations can be converted into Boolean or integer calcs. For example, the following calculation: "If sum([Sales]) > $5000 then 'Show' else 'Hide' end" could be rewritten to be faster as "sum([Sales]) > 5000." It's now a Boolean and will evaluate faster when you drop that on your filter shelf.
- Delete unused calculations in a workbook.
    - Often you will create calculations that you end up not needing. Clean these up before you publish.
- Create ZN and Ifnull calcs to eliminate nulls, which slow things down.
- Eliminate text heavy tables in Tableau. Tableau was built to render graphics first and textual elements second. The more you can visually present information and minimize text, especially large text tables, the faster your workbook will run.
- Avoid using Custom SQL wherever possible. Use Tableau's multiple table and joining options in the data connection window.
    - When adding a table to a data source with joins, select the 'tables' option from the data. Tableau will skip several validation steps in this case.
- Use as few data sources as needed to achieve your analysis. Remove any unused data sources.

You can find more recommendations here:

- http://www.tableau.com/sites/default/files/basic_performance_tips.pdf
- https://www.interworks.com/blog/bfair/2015/02/23/tableau-performance-checklist

Congratulations on finishing the course!

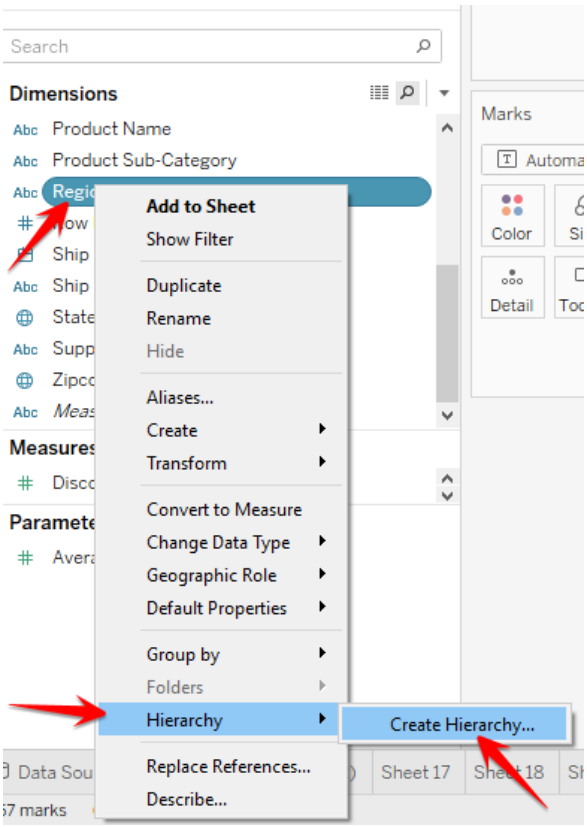
Appendix 1: New Tableau Features

Set Actions

Set Actions are another way that you can improve interactivity on your dashboards. Just as you would use a set to display a summarized version or subset of your data, you can use the Set Actions feature to easily control dashboard interactivity across multiple views using a single click or selection. In this example, we will learn how to use Set Actions to drill into different levels of hierarchy in our data. We will use a hierarchy involving Region, State and City in this tutorial

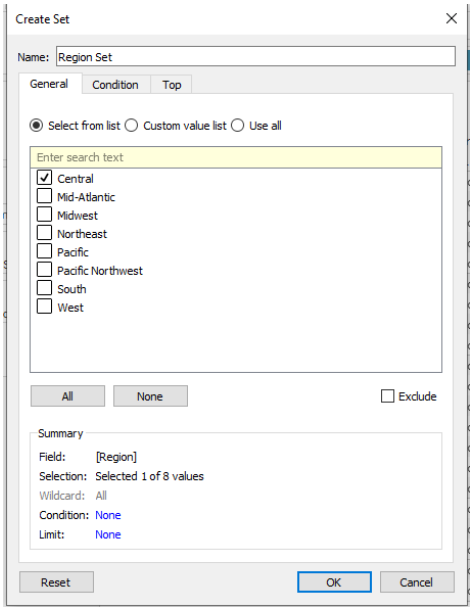
Steps:

1. In your data pane window, right-click on Region > Hierarchy > Create Hierarchy. Name this Hierarchy.



2. Drag and drop State and City into the hierarchy underneath Region.

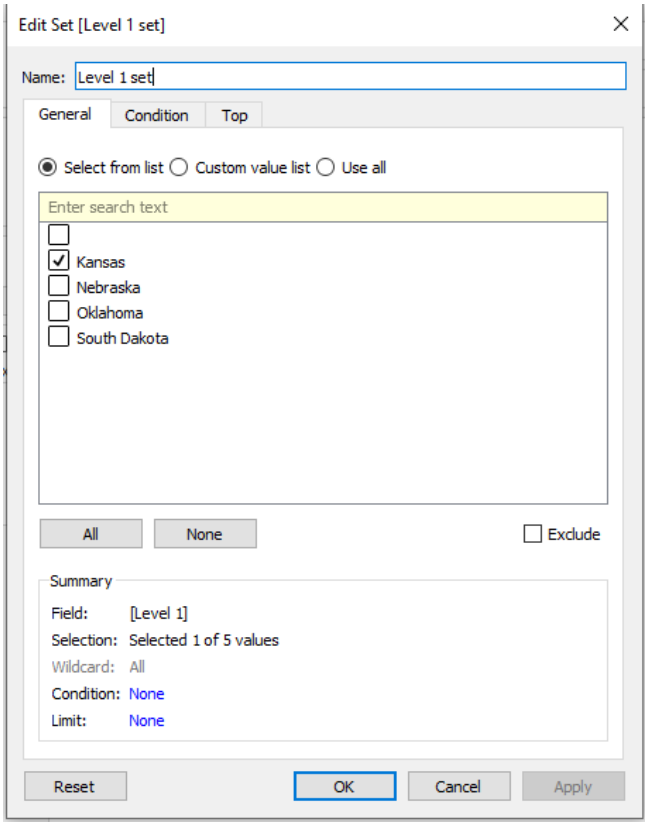
3. Right-click the Region field and create a set. Select any value in the list



4. Name this the “Region Set”.

5. Create a calculation to allow enable us to select the first level in our hierarchy. In the calculation window, create the formula below and name it Level 1. Hit OK.

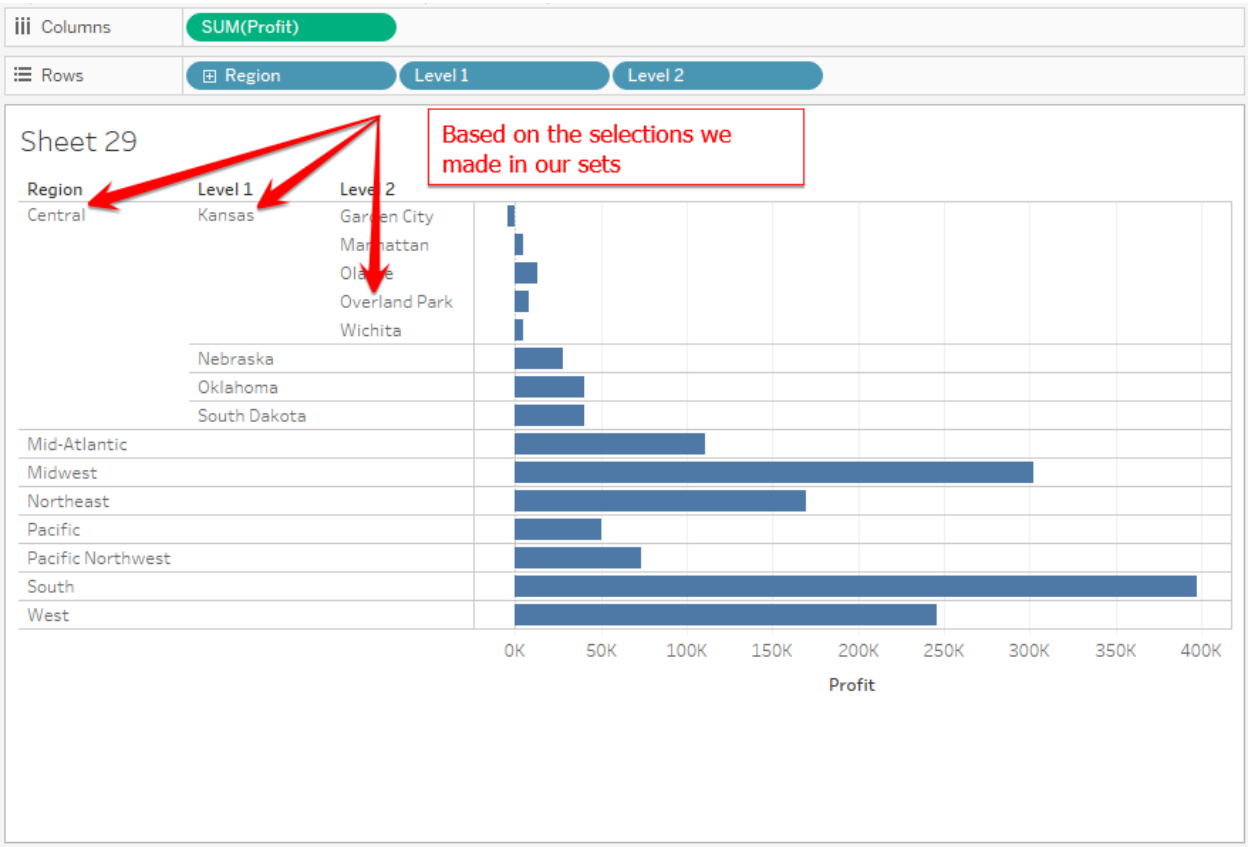
6. Right-click on the calculated field you just created (e.g. Level 1) and create a set. Name this Level 1 set and select any value in that list.



7. Create a calculated field for the second level of the hierarchy, as shown in the screenshot below. Name this Level 2 and hit OK.

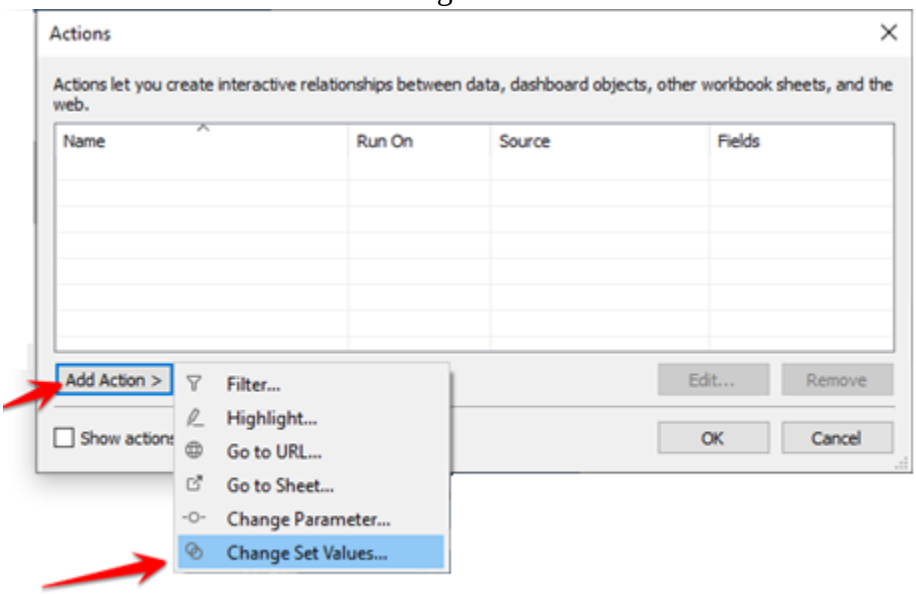
Now we will build a chart using the Level 1 and Level 2 of our hierarchy

8. Drag Region, Level 1, and Level 2 calculations to Rows and Profit to the Columns shelf

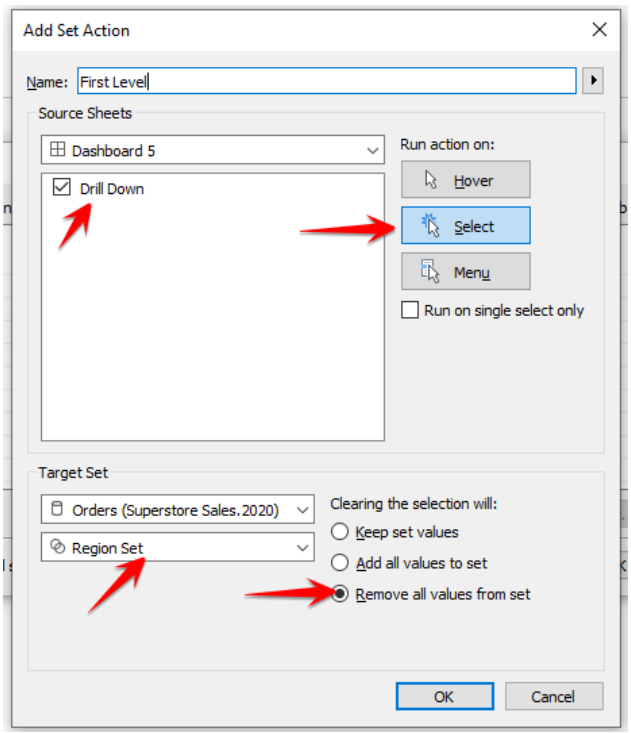


9. Name this chart Drill Down
10. Create a new dashboard and drag the Drill down chart to it.

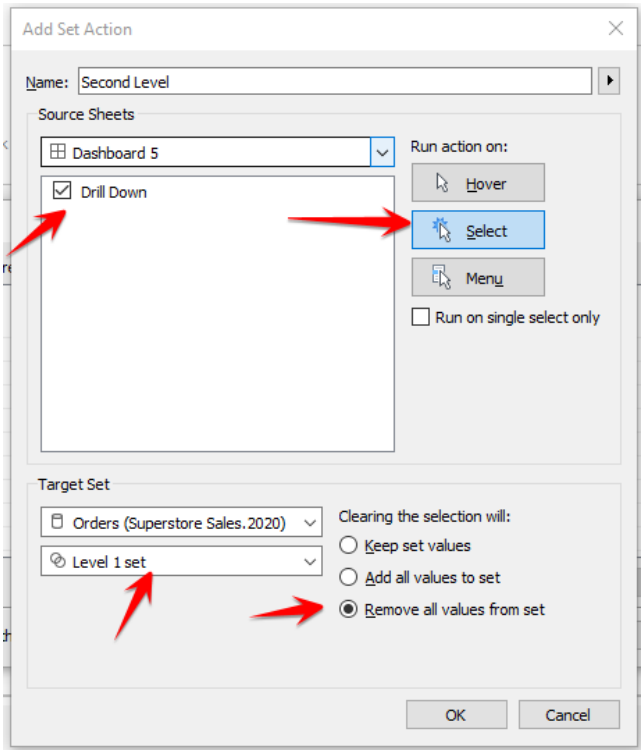
11. Select Dashboard Actions > Change Set Values



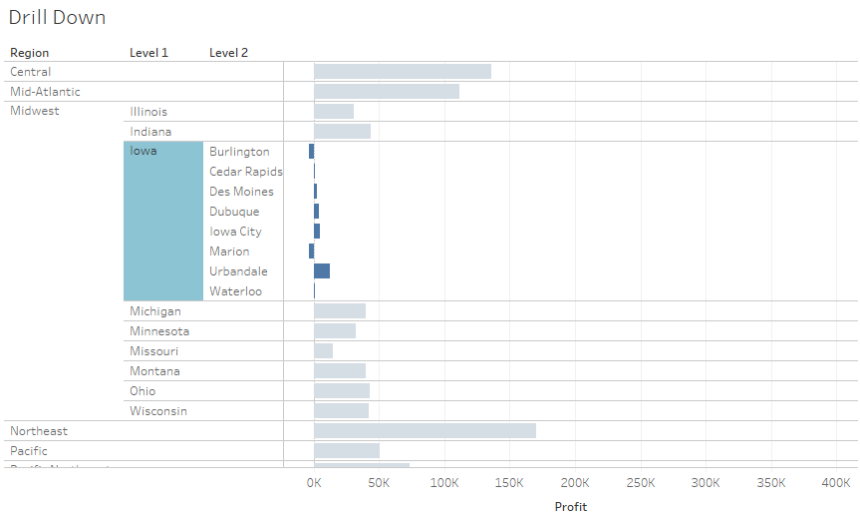
12. Configure the Set Action as shown in the screenshot below for Level 1 of the hierarchy. Name this action First Level.



13. Create a new set action as shown in the screenshot below. Name this Second Level.



14. You can now test the drill down set actions by clicking through the Region, Level 1 and Level 2 on the dashboard.



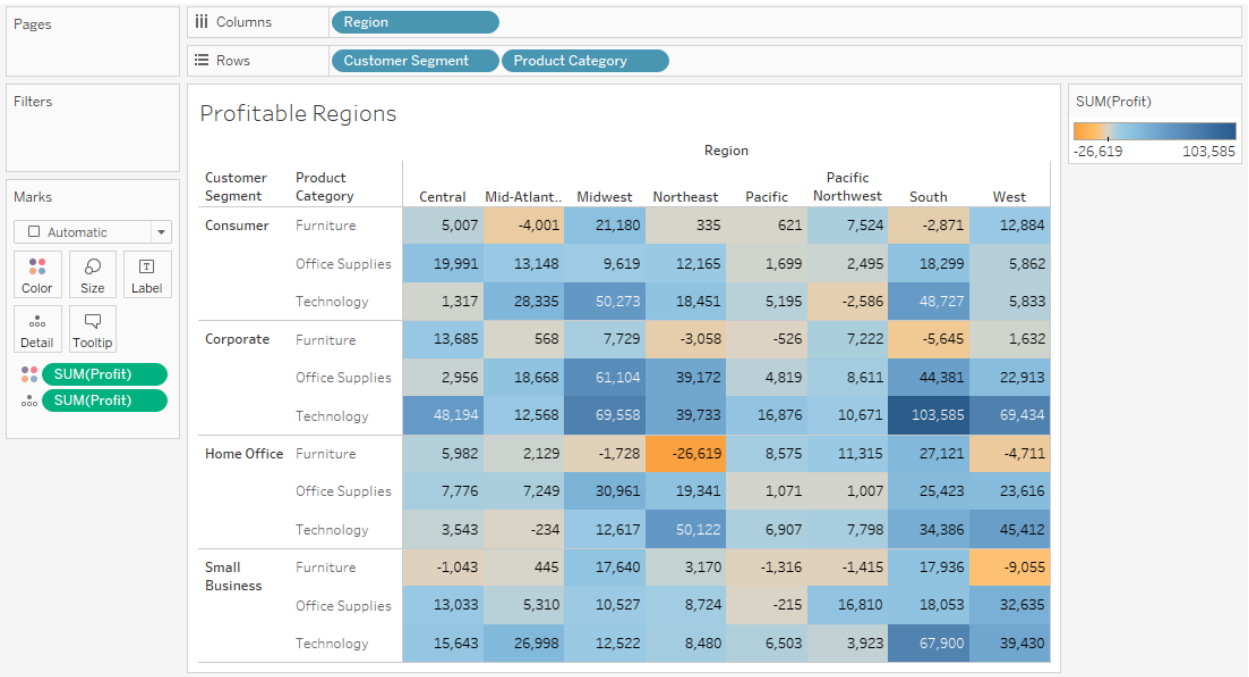
Parameter Actions

Parameter Actions are a new feature in Tableau that provides you the ability to change a parameter value when you directly interact with you chart. In this example, we will look at how to use a parameter action to summarize a subset of the data to show the Average profit from selections we make.

To create:

1. Drag Region to the Column shelf and then place Customer Segment and Product Category to the Row shelf
2. Drag Profit to the color shelf to create a highlight table. Turn on your marks label to display the profit numbers.
3. Drag Profit to the details shelf. This is a crucial step to enable us to interact with this metric when we use the parameter action.

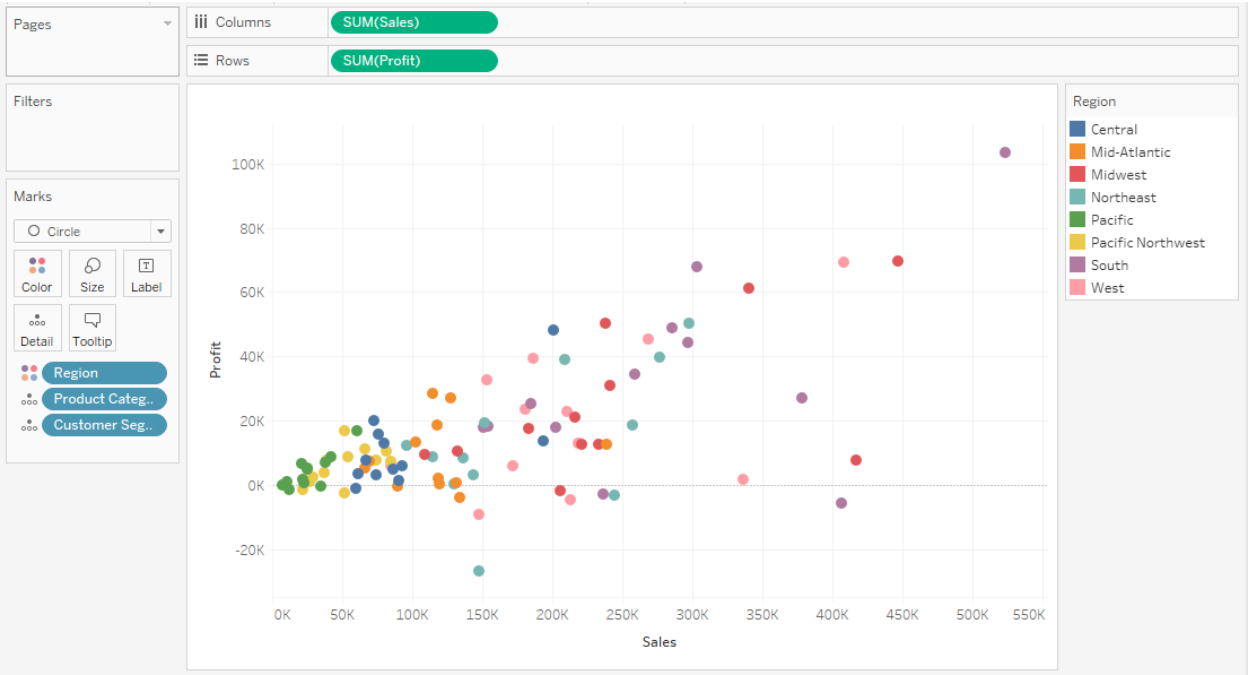
Your chart should look like this:



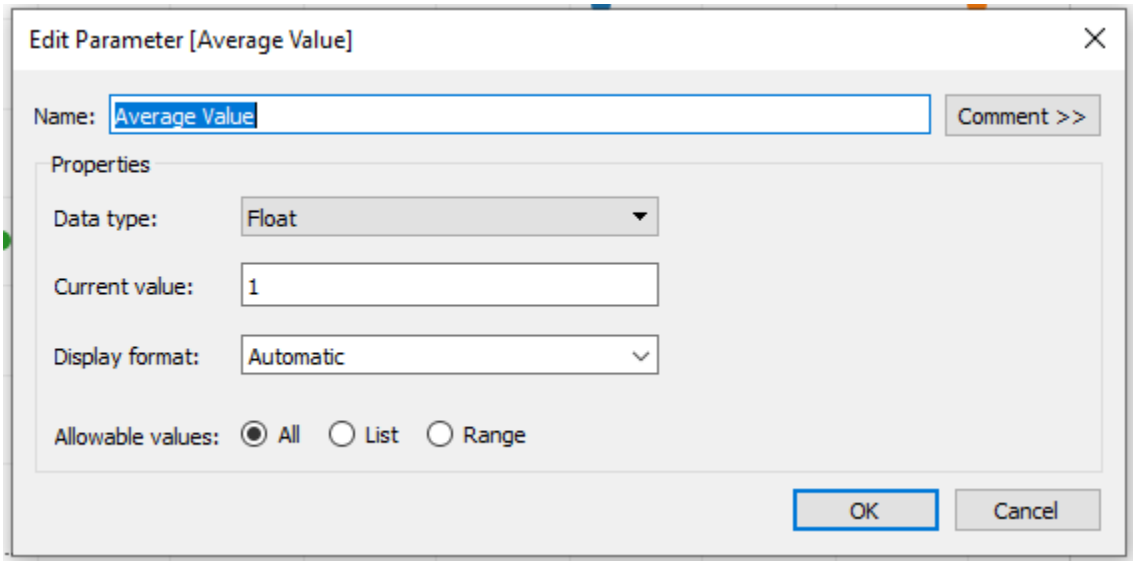
4. Name this chart Profitable Regions

Now let us build a scatterplot using Sales and Profit.

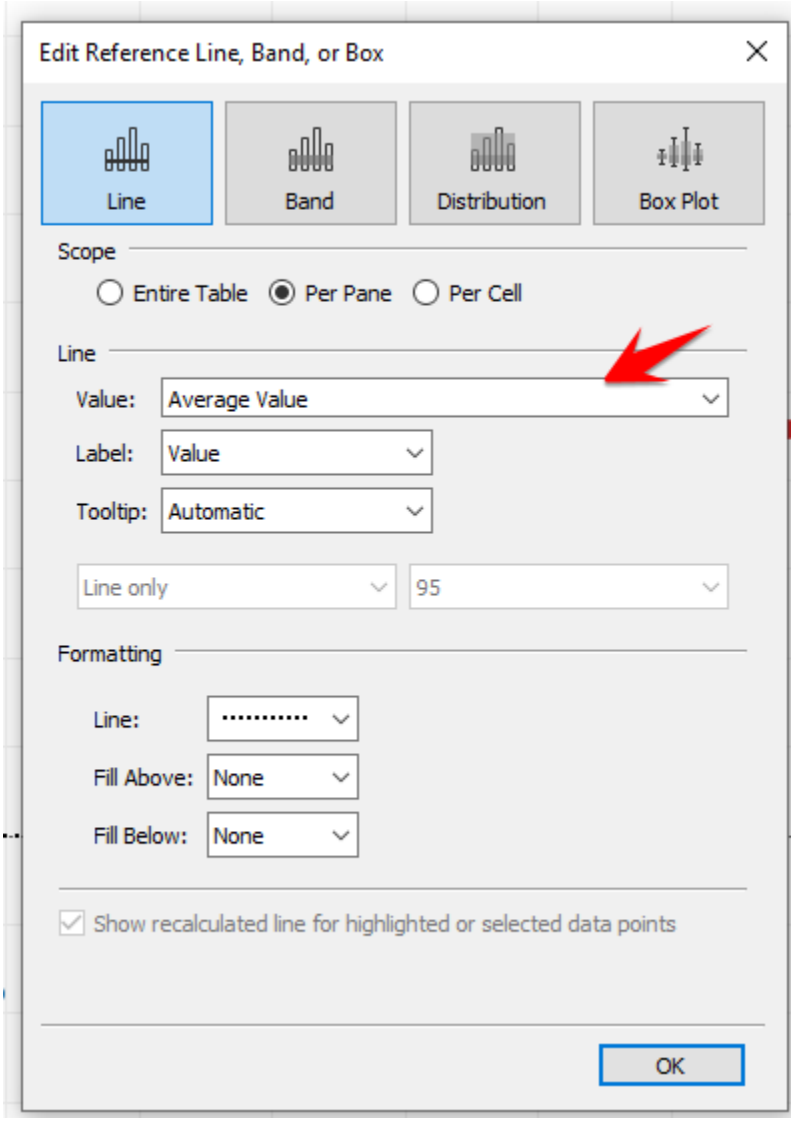
5. Place Customer Segment and Product Category on the detail shelf and Region on color shelf, as shown in the screenshot below:



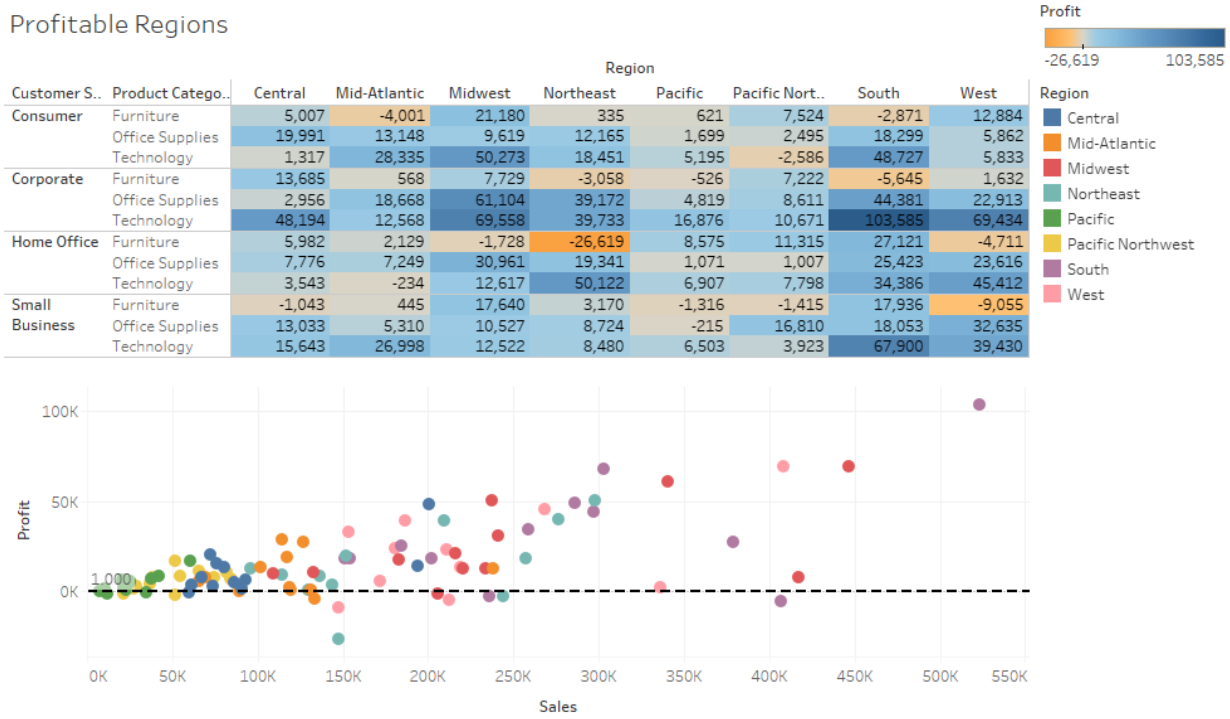
6. Name this chart Scatterplot.
7. Create a parameter and label it Average Value, as shown below:



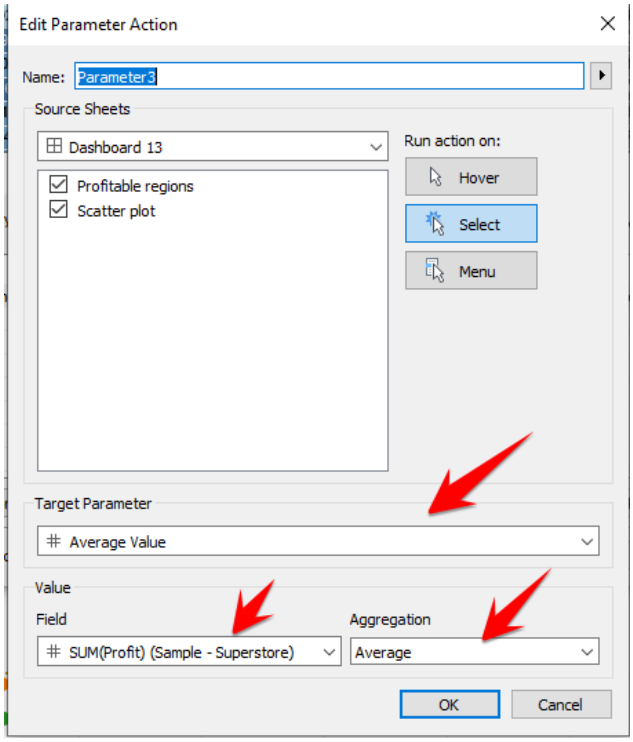
8. On your scatter plot view, add a reference line using the Average Value parameter you just built



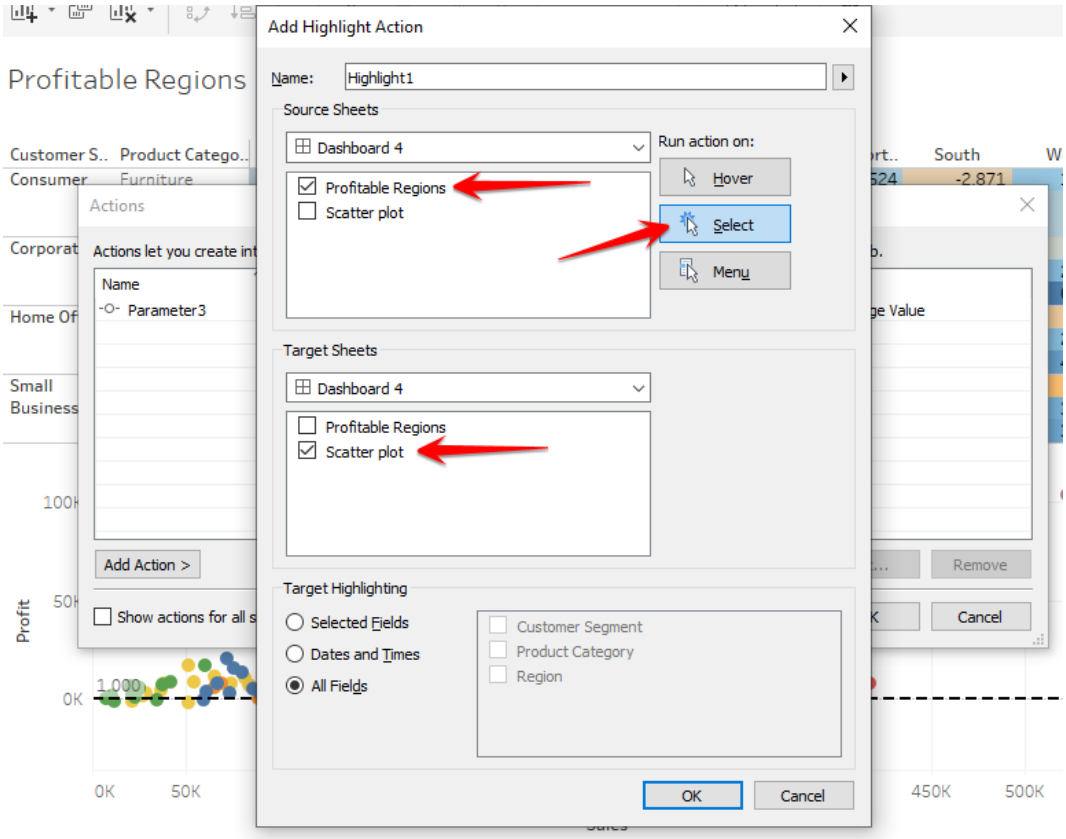
9. Create a new dashboard and bring the two charts you built into the dashboard.



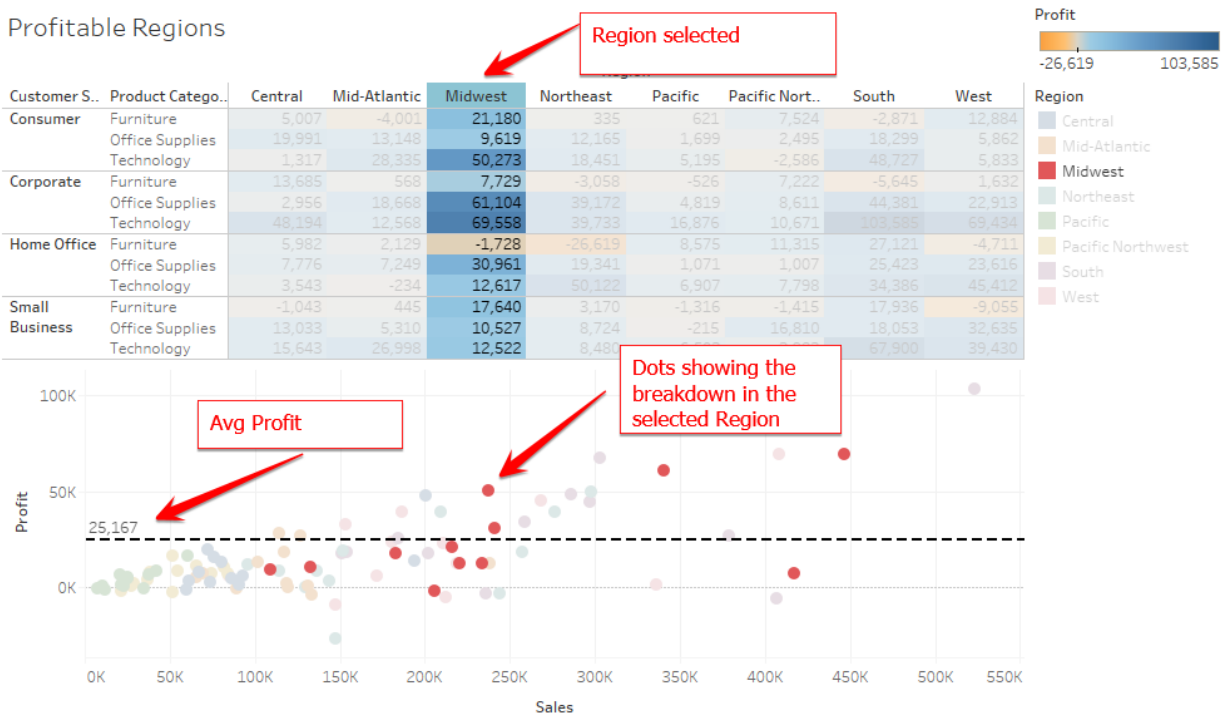
10. From the dashboard menu select Dashboard > Actions > Add Action > Change Parameter.
11. Select the options as shown below and hit OK.

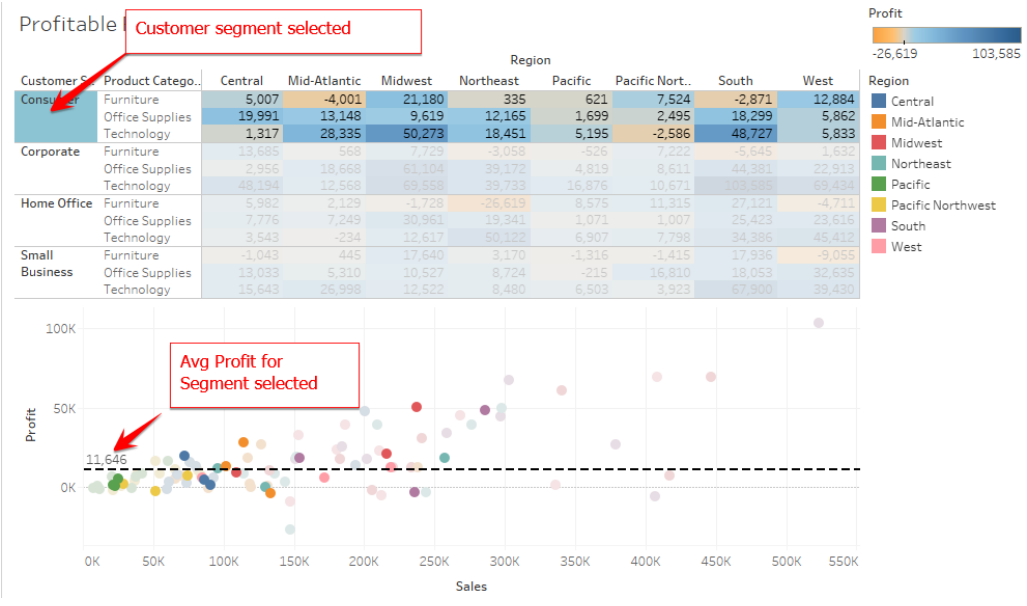


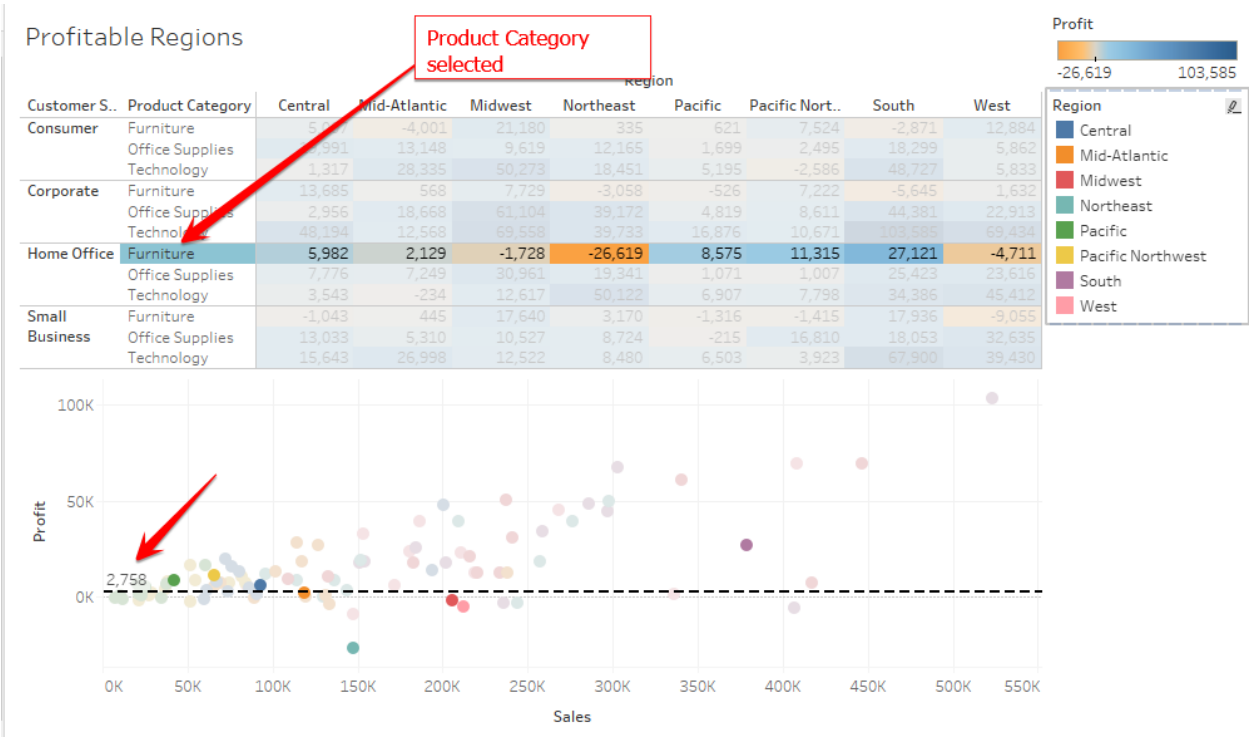
12. In the Actions window select Add Action > Highlight and select the options shown below to highlight marks in the scatter plot

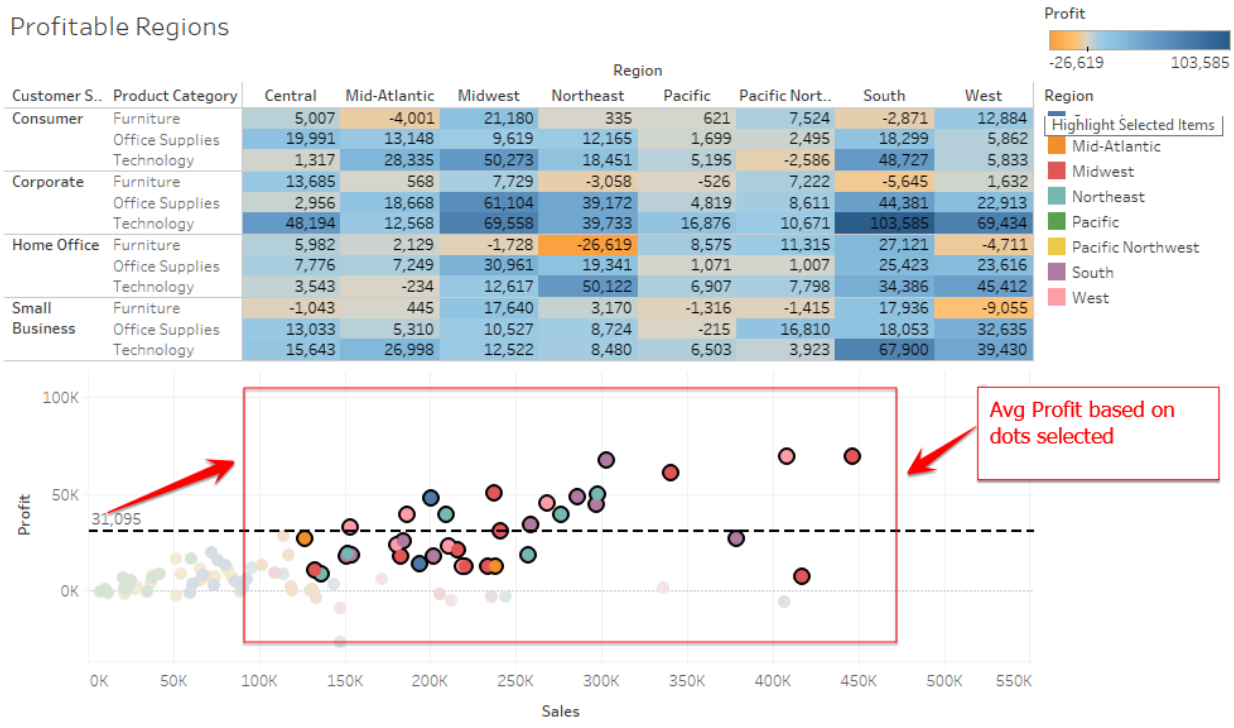


13. Click through the Region, Customer Segment and Product Category in your highlight table to see the reference line update to show you the Average Profit. Additionally, you can select marks in your scatter plot to see the Average Profit value update

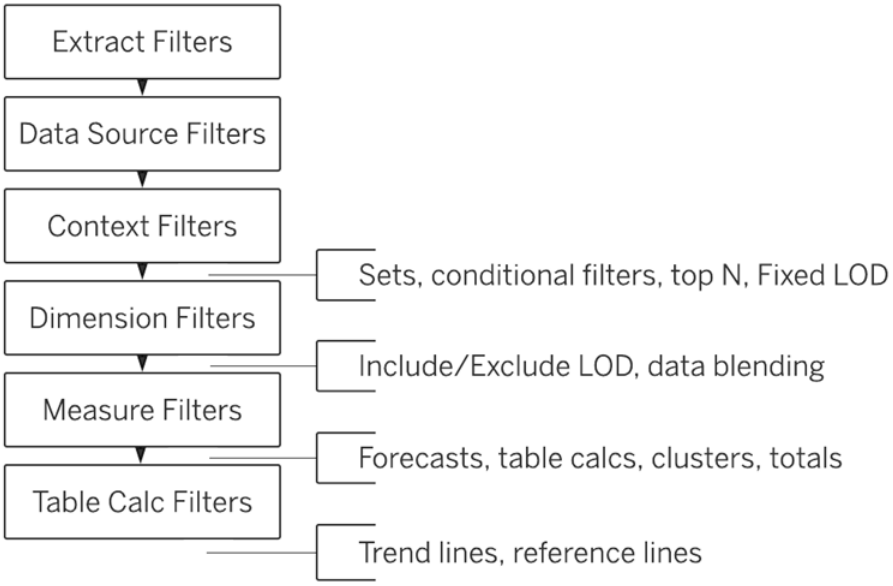

Profitable Regions
Region selected
Profit
-26,619
103,585
Region
Central
Mid-Atlantic
Midwest
Northeast
Pacific
Pacific Northwest
South
West
Customer S..
Product Catego..
Central
Mid-Atlantic
Midwest
Northeast
Pacific
Pacific Nort..
South
West
Consumer
Furniture 5,007 -4,001 21,180 335 621 7,524 -2,871 12,884
Office Supplies 19,991 13,148 9,619 12,165 1,699 2,495 18,299 5,862
Technology 1,317 28,335 50,273 18,451 5,195 -2,586 48,727 5,833
Corporate
Furniture 13,685 568 7,729 -3,058 -526 7,222 -5,645 1,632
Office Supplies 2,956 18,668 61,104 39,172 4,819 8,611 44,381 22,913
Technology 48,194 12,568 69,558 39,733 16,876 10,671 103,585 69,434
Home Office
Furniture 5,982 2,129 -1,728 -26,619 8,575 11,315 27,121 -4,711
Office Supplies 7,776 7,249 30,961 19,341 1,071 1,007 25,423 23,616
Technology 3,543 -234 12,617 50,122 6,907 7,798 34,386 45,412
Small Business
Furniture -1,043 445 17,640 3,170 -1,316 -1,415 17,936 -9,055
Office Supplies 13,033 5,310 10,527 8,724 -215 16,810 18,053 32,635
Technology 15,643 26,998 12,522 8,480 6,503 3,923 67,900 39,430
Dots showing the breakdown in the selected Region
Avg Profit
100K
50K
0K
Profit
25,167
0K
50K
100K
150K
200K
250K
300K
350K
400K
450K
500K
550K
Sales


Profitable
Customer segment selected
Profit
-26,619
103,585
Region
Central
Mid-Atlantic
Midwest
Northeast
Pacific
Pacific Northwest
South
West
Region
Product Catego..
Central
Mid-Atlantic
Midwest
Northeast
Pacific
Pacific Nort..
South
West
Furniture 5,007 -4,001 21,180 335 621 7,524 -2,871 12,884
Office Supplies 19,991 13,148 9,619 12,165 1,699 2,495 18,299 5,862
Technology 1,317 28,335 50,273 18,451 5,195 -2,586 48,727 5,833
Corporate
Furniture 13,685 568 7,729 -3,058 -526 7,222 -5,645 1,632
Office Supplies 2,956 18,668 61,104 39,172 4,819 8,611 44,381 22,913
Technology 48,194 12,568 69,558 39,733 16,876 10,671 103,585 69,434
Home Office
Furniture 5,982 2,129 -1,728 -26,619 8,575 11,315 27,121 -4,711
Office Supplies 7,776 7,249 30,961 19,341 1,071 1,007 25,423 23,616
Technology 3,543 -234 12,617 50,122 6,907 7,798 34,386 45,412
Small Business
Furniture -1,043 445 17,640 3,170 -1,316 -1,415 17,936 -9,055
Office Supplies 13,033 5,310 10,527 8,724 -215 16,810 18,053 32,635
Technology 15,643 26,998 12,522 8,480 6,503 3,923 67,900 39,430
Avg Profit for Segment selected
100K
50K
0K
Profit
11,646
0K
50K
100K
150K
200K
250K
300K
350K
400K
450K
500K
550K
Sales


Profitable Regions
Product Category selected
Profit
-26,619
103,585
Region
Central
Mid-Atlantic
Midwest
Northeast
Pacific
Pacific Northwest
South
West
Region
Customer S..
Product Category
Central
Mid-Atlantic
Midwest
Northeast
Pacific
Pacific Nort..
South
West
Consumer
Furniture
Office Supplies
Technology
Corporate
Furniture
Office Supplies
Technology
Home Office
Furniture
5,982
2,129
-1,728
-26,619
8,575
11,315
27,121
-4,711
Office Supplies
Technology
Small Business
Furniture
Office Supplies
Technology
100K
50K
0K
Profit
2,758
0K
50K
100K
150K
200K
250K
300K
350K
400K
450K
500K
550K
Sales


Profitable Regions
Profit
-26,619
103,585
Region
Highlight Selected Items
Mid-Atlantic
Midwest
Northeast
Pacific
Pacific Northwest
South
West
Customer S..
Product Category
Central
Mid-Atlantic
Midwest
Northeast
Pacific
Pacific Nort..
South
West
Consumer
Furniture
5,007
-4,001
21,180
335
621
7,524
-2,871
12,884
Office Supplies
19,991
13,148
9,619
12,165
1,699
2,495
18,299
5,862
Technology
1,317
28,335
50,273
18,451
5,195
-2,586
48,727
5,833
Corporate
Furniture
13,685
568
7,729
-3,058
-526
7,222
-5,645
1,632
Office Supplies
2,956
18,668
61,104
39,172
4,819
8,611
44,381
22,913
Technology
48,194
12,568
69,558
39,733
16,876
10,671
103,585
69,434
Home Office
Furniture
5,982
2,129
-1,728
-26,619
8,575
11,315
27,121
-4,711
Office Supplies
7,776
7,249
30,961
19,341
1,071
1,007
25,423
23,616
Technology
3,543
-234
12,617
50,122
6,907
7,798
34,386
45,412
Small Business
Furniture
-1,043
445
17,640
3,170
-1,316
-1,415
17,936
-9,055
Office Supplies
13,033
5,310
10,527
8,724
-215
16,810
18,053
32,635
Technology
15,643
26,998
12,522
8,480
6,503
3,923
67,900
39,430
100K
50K
0K
Profit
31,095
0K
50K
100K
150K
200K
250K
300K
350K
400K
450K
500K
550K
Sales
Avg Profit based on dots selected

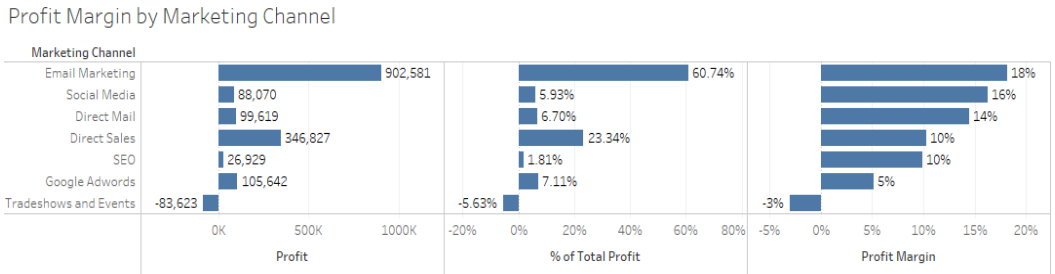
Appendix 2: Tableau's Filter Order of Operations



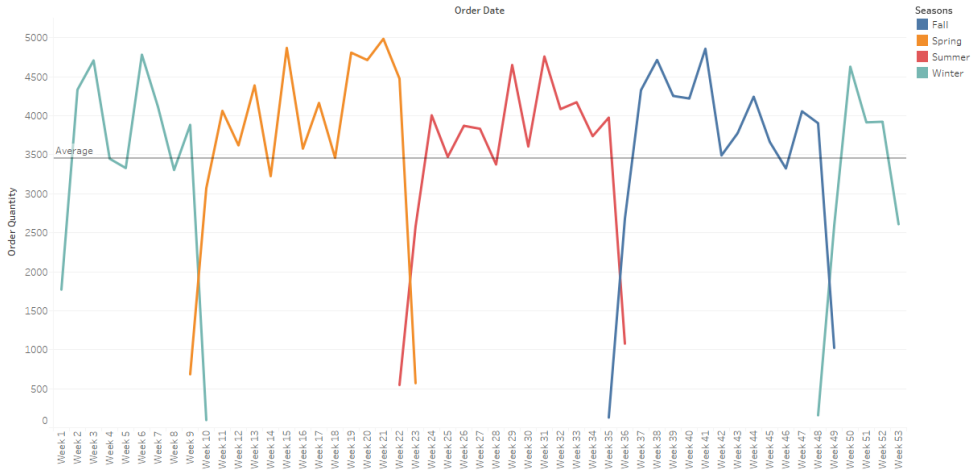
Appendix 3: Solutions to Practice Exercises

Note that the answers below are possible solutions, but not the only solution. Since Tableau is infinitely customizable, you may end with a chart in different colors, shapes, size, or in some instances, even a different chart type, especially as you advance to more advanced topics. If the same conclusions are easily identifiable in your chart or dashboard, it may still be an appropriate solution, even though some aspects look different.

Pg. 13

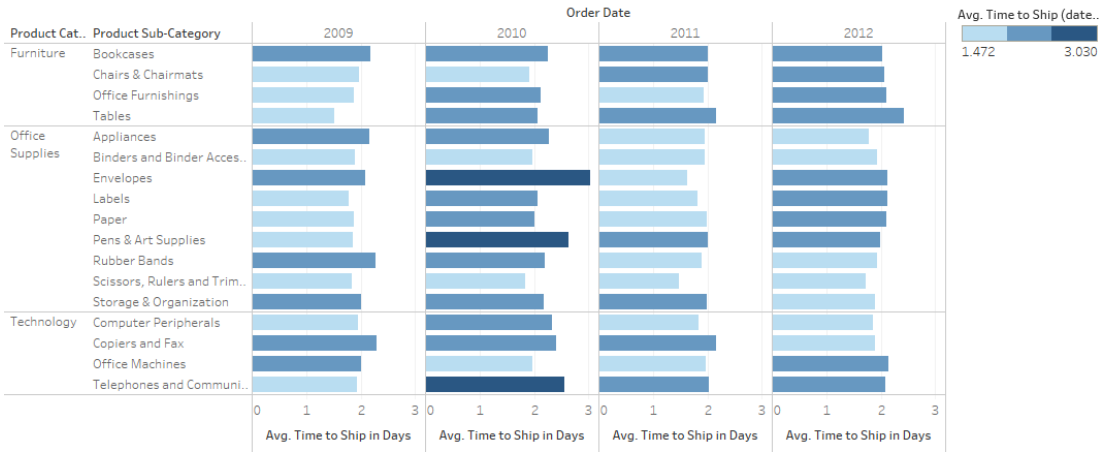


Order Quantity by Day

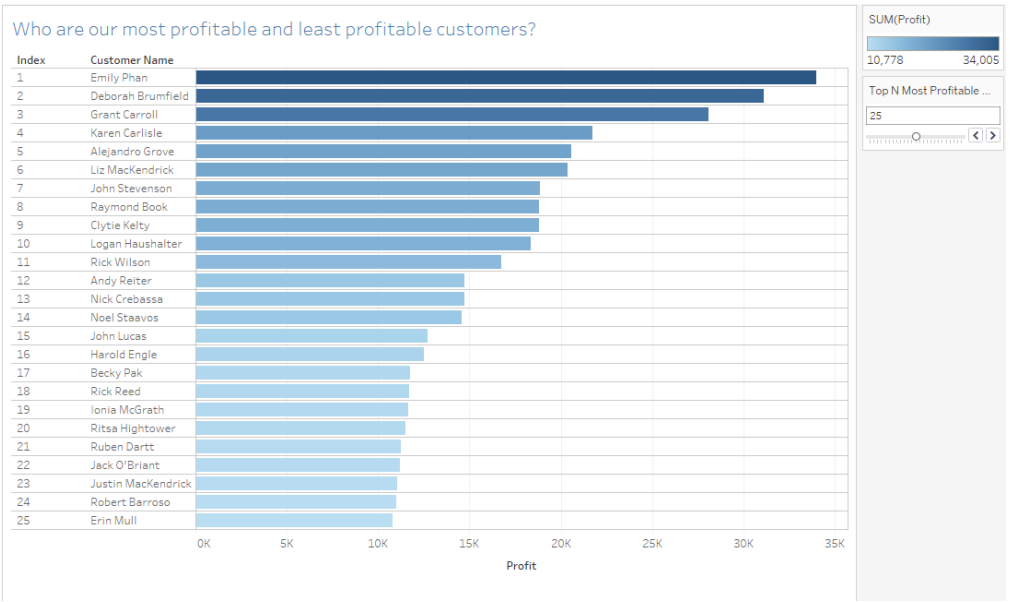


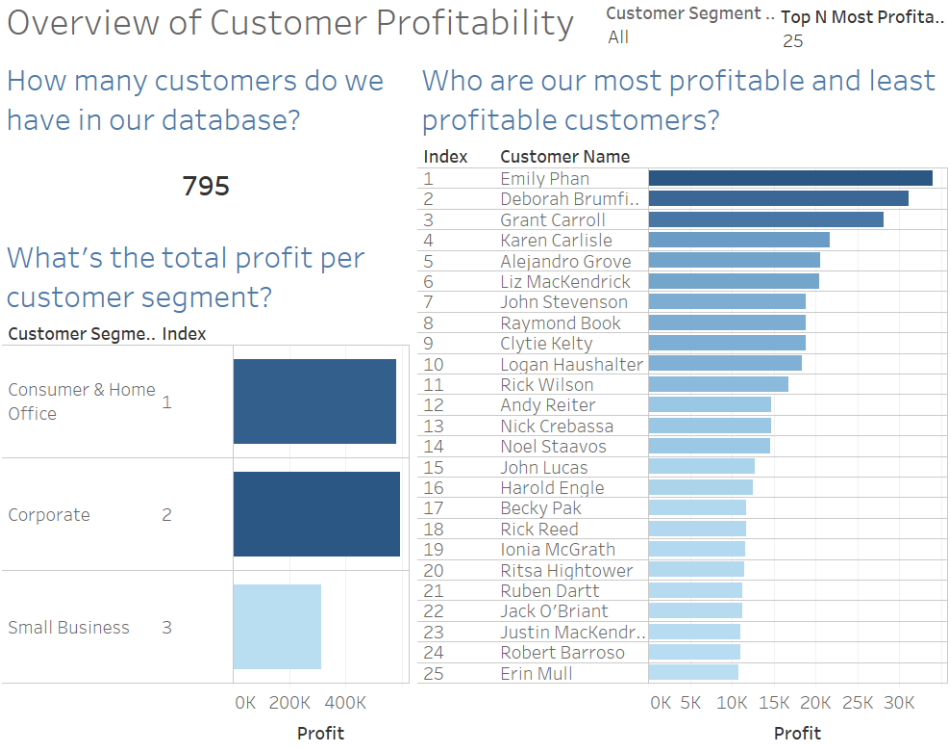
Pg. 18

Time to Ship

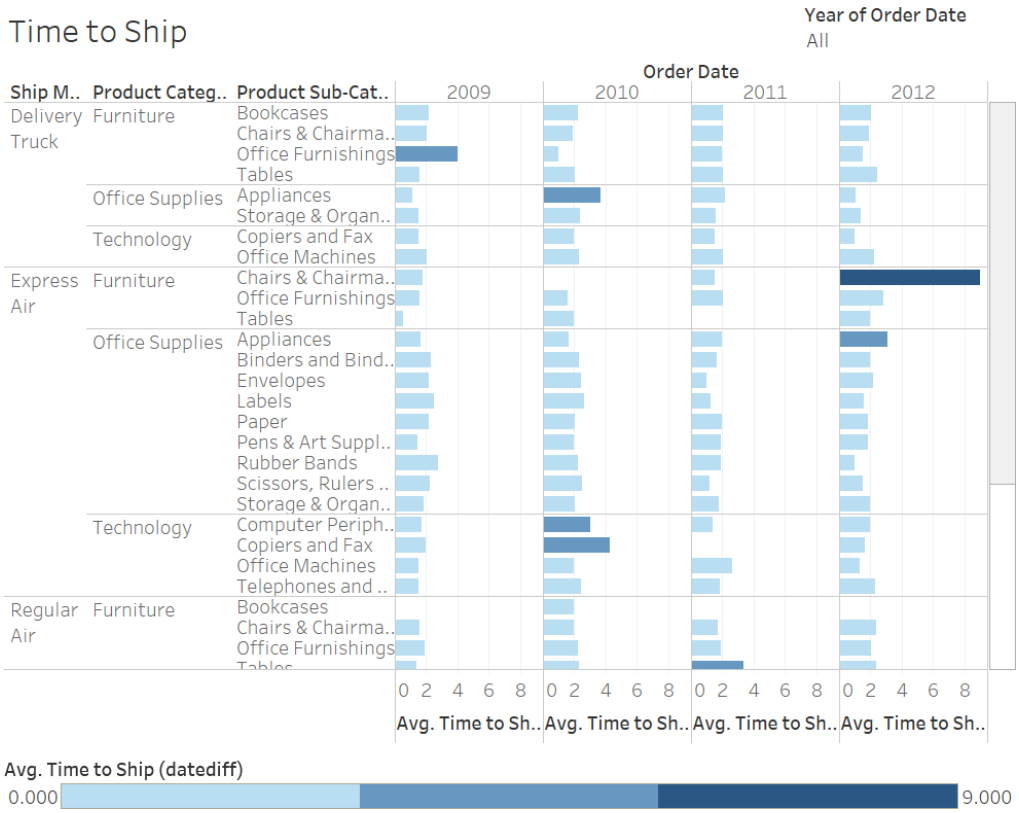


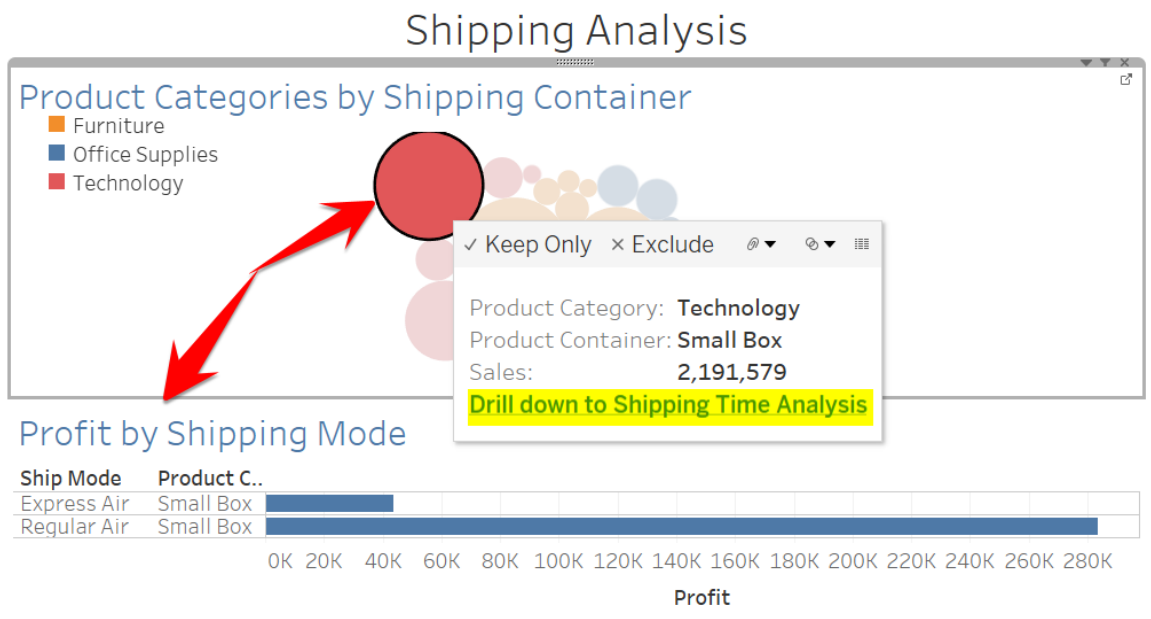
Pg. 28

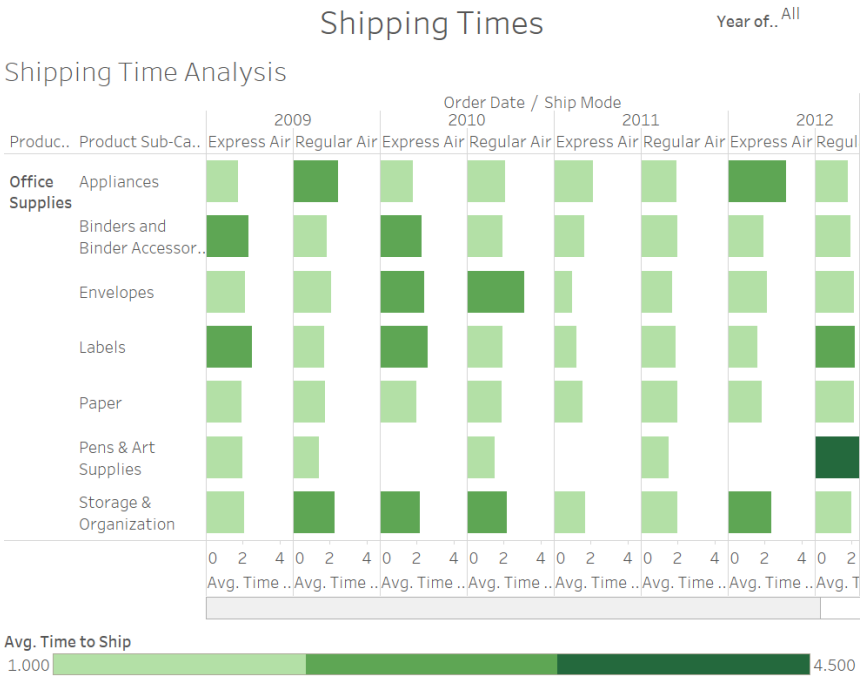

Who are our most profitable and least profitable customers?
Index
Customer Name
1 Emily Phan
2 Deborah Brumfield
3 Grant Carroll
4 Karen Carlisle
5 Alejandro Grove
6 Liz MacKendrick
7 John Stevenson
8 Raymond Book
9 Clytie Kelty
10 Logan Haushalter
11 Rick Wilson
12 Andy Reiter
13 Nick Crebassa
14 Noel Staavos
15 John Lucas
16 Harold Engle
17 Becky Pak
18 Rick Reed
19 Ionia McGrath
20 Ritsa Hightower
21 Ruben Dartt
22 Jack O'Briant
23 Justin MacKendrick
24 Robert Barroso
25 Erin Mull
0K 5K 10K 15K 20K 25K 30K 35K
Profit
SUM(Profit)
10,778 34,005
Top N Most Profitable ...
25


Overview of Customer Profitability
Customer Segment ..
All
Top N Most Profita..
25
How many customers do we have in our database?
795
What's the total profit per customer segment?
Customer Segme.. Index
Consumer & Home Office 1
Corporate 2
Small Business 3
0K 200K 400K
Profit
Who are our most profitable and least profitable customers?
Index
Customer Name
1 Emily Phan
2 Deborah Brumfi..
3 Grant Carroll
4 Karen Carlisle
5 Alejandro Grove
6 Liz MacKendrick
7 John Stevenson
8 Raymond Book
9 Clytie Kelty
10 Logan Haushalter
11 Rick Wilson
12 Andy Reiter
13 Nick Crebassa
14 Noel Staavos
15 John Lucas
16 Harold Engle
17 Becky Pak
18 Rick Reed
19 Ionia McGrath
20 Ritsa Hightower
21 Ruben Dartt
22 Jack O'Briant
23 Justin MacKendr..
24 Robert Barroso
25 Erin Mull
0K 5K 10K 15K 20K 25K 30K
Profit

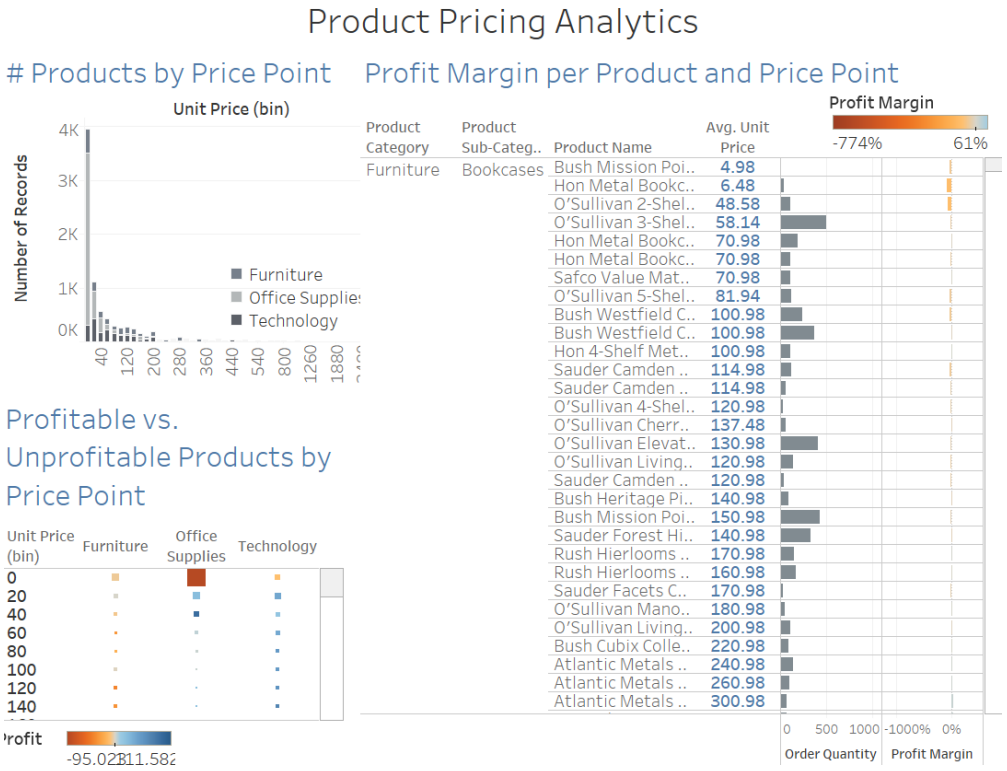
Pg. 33


Time to Ship
Year of Order Date
All
Order Date
Ship M..
Product Categ..
Product Sub-Cat..
2009
2010
2011
2012
Delivery Truck
Furniture
Bookcases
Chairs & Chairma..
Office Furnishings
Tables
Office Supplies
Appliances
Storage & Organ..
Technology
Copiers and Fax
Office Machines
Express Air
Furniture
Chairs & Chairma..
Office Furnishings
Tables
Office Supplies
Appliances
Binders and Bind..
Envelopes
Labels
Paper
Pens & Art Suppl..
Rubber Bands
Scissors, Rulers ..
Storage & Organ..
Technology
Computer Periph..
Copiers and Fax
Office Machines
Telephones and ..
Regular Air
Furniture
Bookcases
Chairs & Chairma..
Office Furnishings
Tables
0 2 4 6 8
Avg. Time to Sh..
Avg. Time to Ship (datediff)
0.000
9.000

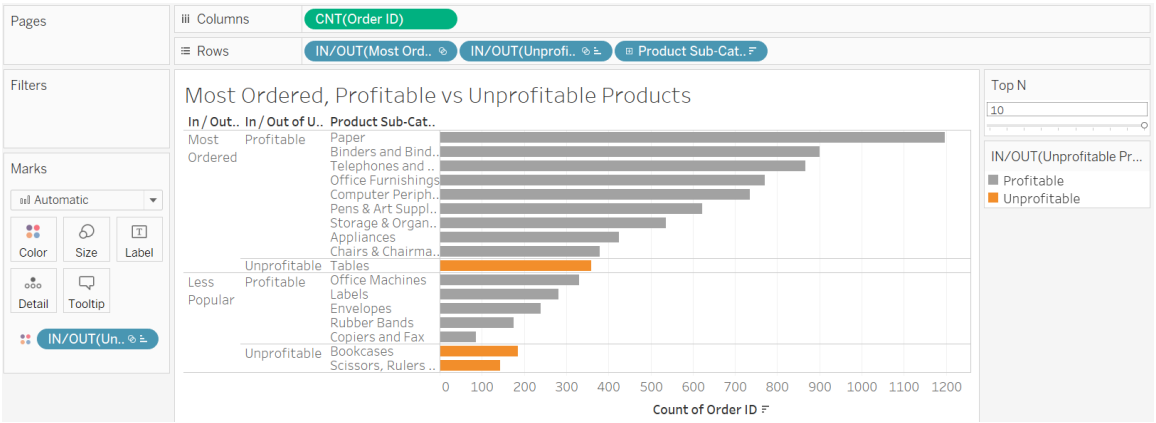

Shipping Analysis
Product Categories by Shipping Container
Furniture
Office Supplies
Technology
✓ Keep Only
× Exclude
Product Category: Technology
Product Container: Small Box
Sales: 2,191,579
Drill down to Shipping Time Analysis
Profit by Shipping Mode
Ship Mode
Product C..
Express Air
Small Box
Regular Air
Small Box
0K 20K 40K 60K 80K 100K 120K 140K 160K 180K 200K 220K 240K 260K 280K
Profit


Shipping Times
Year of.. All
Shipping Time Analysis
Order Date / Ship Mode
2009
2010
2011
2012
Produc..
Product Sub-Ca..
Express Air
Regular Air
Office Supplies
Appliances
Binders and Binder Accessor..
Envelopes
Labels
Paper
Pens & Art Supplies
Storage & Organization
Avg. Time ..
Avg. Time to Ship
1.000
4.500

Pg. 45

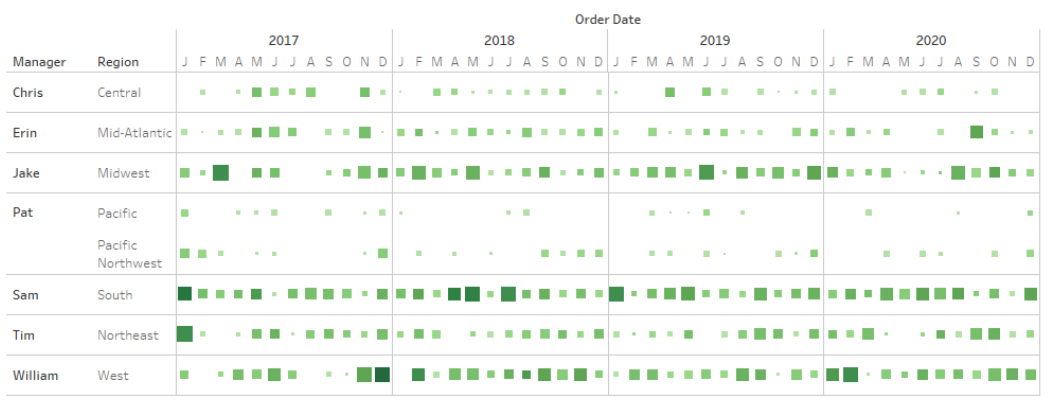

Product Pricing Analytics
Products by Price Point
Unit Price (bin)
Number of Records
Furniture
Office Supplies
Technology
Profitable vs. Unprofitable Products by Price Point
Unit Price (bin)
Furniture
Office Supplies
Technology
Profit
Profit Margin and Price Point
Profit Margin
-774%
61%
Product Category
Product Sub-Categ..
Product Name
Avg. Unit Price
Furniture
Bookcases
Bush Mission Poi.. 4.98
Hon Metal Bookc.. 6.48
O'Sullivan 2-Shel.. 48.58
O'Sullivan 3-Shel.. 58.14
Hon Metal Bookc.. 70.98
Hon Metal Bookc.. 70.98
Safco Value Mat.. 70.98
O'Sullivan 5-Shel.. 81.94
Bush Westfield C.. 100.98
Bush Westfield C.. 100.98
Hon 4-Shelf Met.. 100.98
Sauder Camden .. 114.98
Sauder Camden .. 114.98
O'Sullivan 4-Shel.. 120.98
O'Sullivan Cherr.. 137.48
O'Sullivan Elevat.. 130.98
O'Sullivan Living.. 120.98
Sauder Camden .. 120.98
Bush Heritage Pi.. 140.98
Bush Mission Poi.. 150.98
Sauder Forest Hi.. 140.98
Rush Hierlooms .. 170.98
Rush Hierlooms .. 160.98
Sauder Facets C.. 170.98
O'Sullivan Mano.. 180.98
O'Sullivan Living.. 200.98
Bush Cubix Colle.. 220.98
Atlantic Metals .. 240.98
Atlantic Metals .. 260.98
Atlantic Metals .. 300.98
Order Quantity
Profit Margin

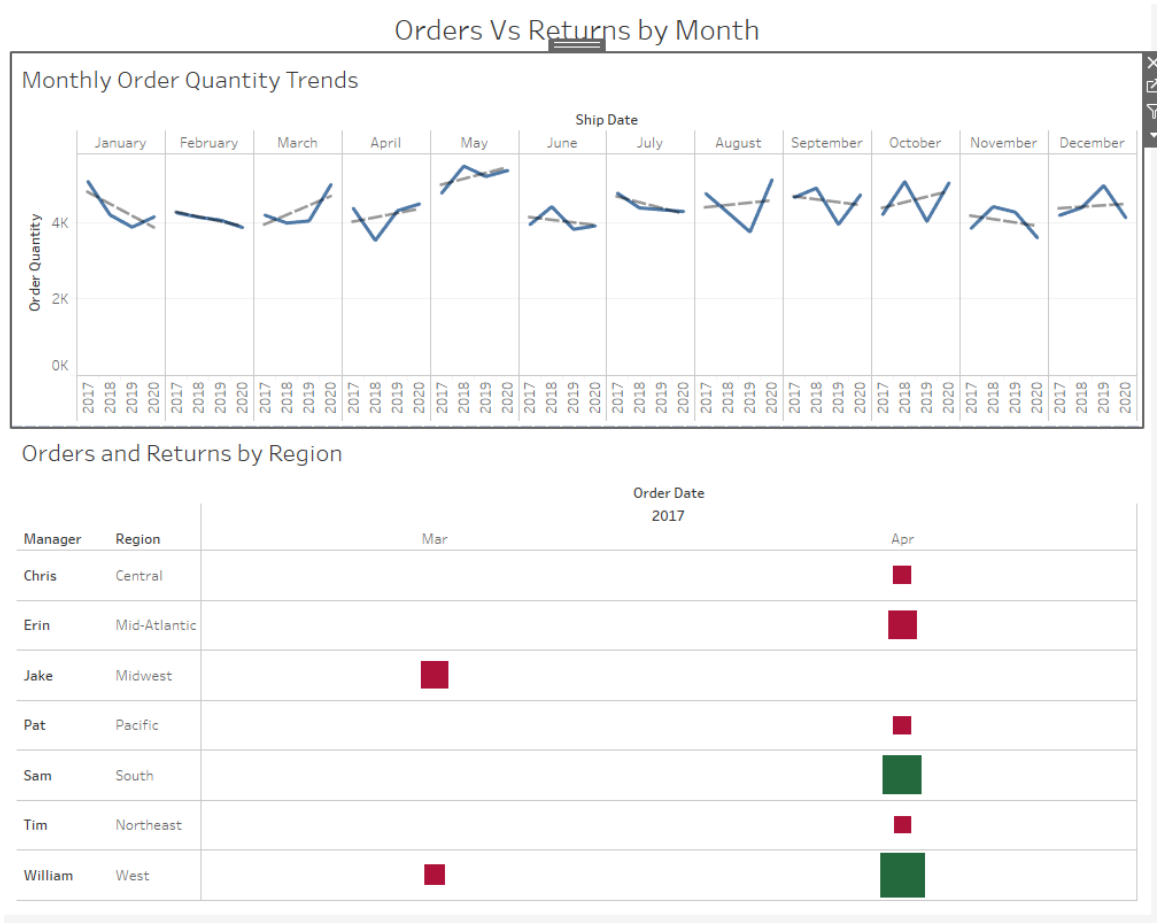
Pg. 56

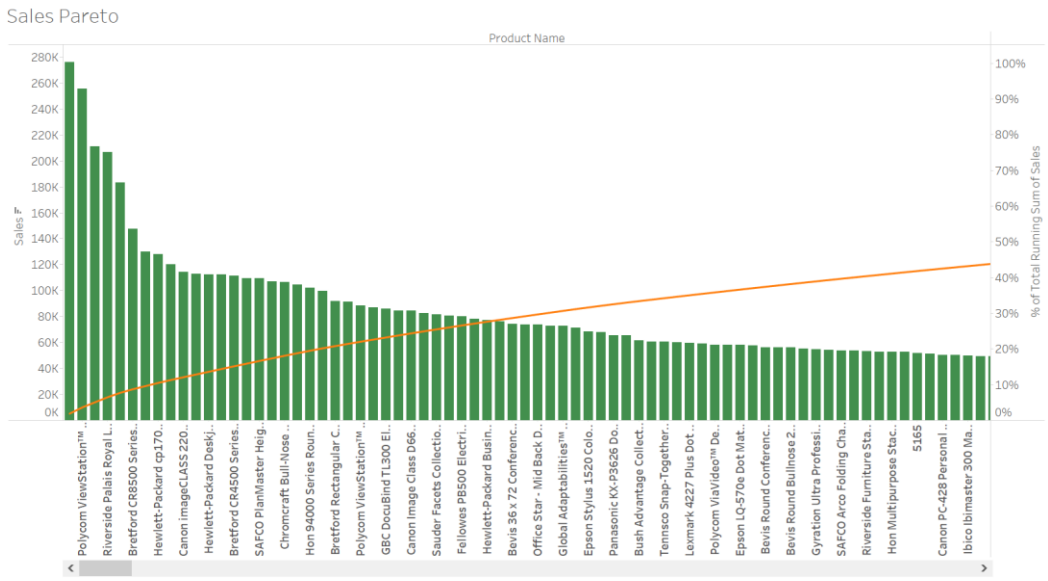

Pages
Columns
CNT(Order ID)
Rows
IN/OUT(Most Ord..
IN/OUT(Unprofi..
Product Sub-Cat..
Filters
Marks
Automatic
Color
Size
Label
Detail
Tooltip
IN/OUT(Un..
Most Ordered, Profitable vs Unprofitable Products
In / Out..
In / Out of U..
Product Sub-Cat..
Most Ordered
Profitable
Paper
Binders and Bind..
Telephones and ..
Office Furnishings
Computer Periph..
Pens & Art Suppl..
Storage & Organ..
Appliances
Chairs & Chairma..
Unprofitable
Tables
Less Popular
Profitable
Office Machines
Labels
Envelopes
Rubber Bands
Copiers and Fax
Unprofitable
Bookcases
Scissors, Rulers ..
0 100 200 300 400 500 600 700 800 900 1000 1100 1200
Count of Order ID
Top N
10
IN/OUT(Unprofitable Pr...
Profitable
Unprofitable

Pg. 92

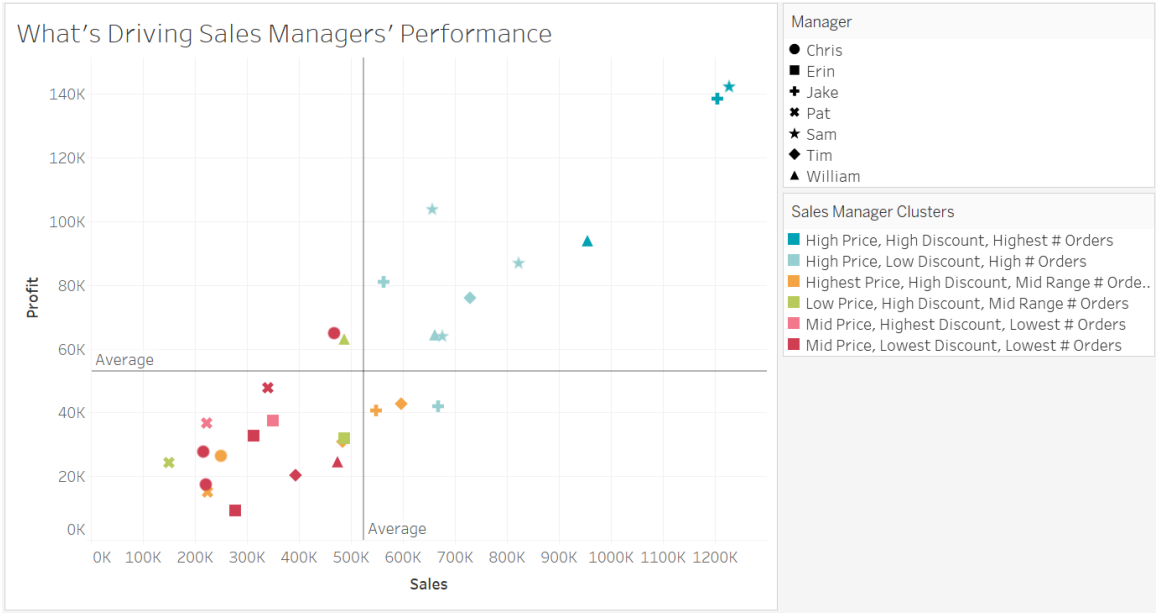

Orders Vs Returns by Month
Monthly Order Quantity Trends
Ship Date
January
February
March
April
May
June
July
August
September
October
November
December
Order Quantity
4K
2K
0K
Keep Only
Exclude
Month of Ship Date: April
Year of Ship Date: 2017
Order Quantity: 4,377
See Returns by Month, Year and Sales Mgr


Orders and Returns by Region
Order Date
2017
2018
2019
2020
Manager
Region
Chris
Central
Erin
Mid-Atlantic
Jake
Midwest
Pat
Pacific
Pacific Northwest
Sam
South
Tim
Northeast
William
West


Orders Vs Returns by Month
Monthly Order Quantity Trends
Ship Date
January
February
March
April
May
June
July
August
September
October
November
December
Order Quantity
4K
2K
0K
2017
2018
2019
2020
Orders and Returns by Region
Order Date
2017
Mar
Apr
Manager
Region
Chris
Central
Erin
Mid-Atlantic
Jake
Midwest
Pat
Pacific
Sam
South
Tim
Northeast
William
West


Sales Pareto
Product Name
Sales
% of Total Running Sum of Sales
280K
260K
240K
220K
200K
180K
160K
140K
120K
100K
80K
60K
40K
20K
0K
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Polycom ViewStation™...
Riverside Palais Royal L...
Bretford CR8500 Series...
Hewlett-Packard cp170...
Canon imageCLASS 220...
Hewlett-Packard DeskJ...
Bretford CR4500 Series...
SAFCO PlanMaster Heig...
Chromcraft Bull-Nose ...
Hon 94000 Series Roun...
Bretford Rectangular C...
Polycom ViewStation™...
GBC DocuBind TL300 El...
Canon Image Class D66...
Sauder Facets Collectio...
Fellowes PB500 Electri...
Hewlett-Packard Busin...
Bevis 36 x 72 Conferenc...
Office Star - Mid Back D...
Global Adaptabilities™...
Epson Stylus 1520 Colo...
Panasonic KX-P3626 Do...
Bush Advantage Collect...
Tennsco Snap-Together...
Lexmark 4227 Plus Dot ...
Polycom ViaVideo™ De...
Epson LQ-570e Dot Mat...
Bevis Round Conferenc...
Bevis Round Bullnose 2...
Gyration Ultra Professi...
SAFCO Arco Folding Cha...
Riverside Furniture Sta...
Hon Multipurpose Stac...
5165
Canon PC-428 Personal ...
Ibico Ibimaster 300 Ma...

Pg. 104


What's Driving Sales Managers' Performance
Profit
Sales
140K
120K
100K
80K
60K
40K
20K
0K
Average
Average
0K 100K 200K 300K 400K 500K 600K 700K 800K 900K 1000K 1100K 1200K
Manager
Chris
Erin
Jake
Pat
Sam
Tim
William
Sales Manager Clusters
High Price, High Discount, Highest # Orders
High Price, Low Discount, High # Orders
Highest Price, High Discount, Mid Range # Orde..
Low Price, High Discount, Mid Range # Orders
Mid Price, Highest Discount, Lowest # Orders
Mid Price, Lowest Discount, Lowest # Orders

Appendix 4: Pop Quiz Answers

Pop Quiz 2

1. Question: If I wanted to compare profit to sales within the same chart, what would be the best chart to use?

   Answer: Bar in bar chart
   Reason: Bar in bars compare two measures along the same scale, so that you can see how much one compares vs the other, or the percent contribution one measure makes vs. another.
2. Question: If I wanted to show length of time for a series of tasks within a project, or show amount of time between two events, what would be the best chart type to use?

   Answer: Gantt chart
   Reason: Gantts are the default project planning or tracking chart and are great for showing steps in a process and how long each will take.
3. Question: If I wanted to see an end result vs. a starting point and eliminate distracting data, would be the best chart to use?

   Answer: Slope chart
   Reason: Slope charts eliminate noise from a time series or trend and show you just where you started and where you finished. It's helpful if you just want to see if you made progress, or where you ended up.
4. Question: If I wanted to see how we compare against a target value, or how close we are to a certain goal, which chart should I use?

   Answer: Bullet graph
   Reason: Bullet graphs are great for showing a target and progress towards that. It encodes a lot of value in a small amount of space, much more so than a gauge like this:

Knowing Tableau makes you valuable to your employer

Demand for people who can analyze business data and provide meaningful and actionable insights are in high demand. The faster you can find insights, the faster you can help drive action that moves your company forward and earns you greater recognition in your work. This book will take you from an intermediate level to advanced, and will have a special focus on adding calculations, interactivity and deeper analysis, and automating reports.

Learn how to:

- Create more complex calculations
- Add more powerful interactivity using parameters and actions
- Build more complex chart types
- Utilize advanced methods to segment your data to mine for critical insights

Ryan Nokes was an early adopter of Tableau and has been using the software for over a decade. This book was born from years of in-the-trenches daily use and will teach you the things you most need to understand. He is a certified Tableau expert and author of numerous books on the software and has consulted or trained from Fortune 500 to small business in myriad industries and departments.