



On-Demand Training: Introduction to Calculations Transcript

Welcome to the Introductions to Calculations video, please feel free to follow along with the companion workbook that's attached to the video because it already contains the instructions and the data sets we'll be using.

Today we're going to go just over the discussion around different types of calculations; and by that I mean Row Level vs. Aggregate vs. Quick Table Calcs, just to give you a general overview of calculations in Tableau. There are additions videos that cover specific function types which can also be found on our training page.

So in this case I'm going to start by making a mistake that most new users of Tableau make. I'm going to explain why they make it, and how you fix it.

I'm going to start with a blank sheet here, and I'm going to bring out Row ID. This is simply a field that has a number that's in sequence here for every single row in our data set, and I'm going to bring out my Profit to 'Text' and I'm going to double-click 'Sales' so that that's put into the table as well. So, what I'm going to do here is create a profit ratio, so I'm just going to right-click my 'Profit' field, choose 'Create Calculated Field', the same dialogue box can be reached by going to 'Analysis' and using 'Create Calculated Field' as well.

So I'm going to take here my profit divided by my sales, I'll call this my 'Profit Ratio Wrong' because I know this is going to return incorrect results; and to make sure that this shows up correctly, I'm going to right-click this field and change the Number Format here to 'Percentage' and I'm going to bring this up my view, you'll say: "Hey, that look's right, what's going on here? Why is that incorrect?" And you can't tell because at this point we're doing a row-level calculation, so for every single row what we are doing here is we are dividing profit by sales and then for the next row we're doing that exact same calculation and so forth.

But really what we want to do here is a different calculation, we want to do our 'Profit Ratio Correct' calculation, which in lieu of doing the set row level, adds in aggregation here – so, the sum of my profit divided by the sum of my sales. Just as a side note here, you'll notice that I didn't put the brackets around sales here but it's still functioning correctly. That's because you don't need the brackets around single-word dimensions or measures, but you will for things like number brackets once there are spaces and multiple words.

I'm going to click 'OK' here and again I'm going to format this as a percentage quickly here so we see the correct information. And I'm going to drag this out, the first thing you'll notice here is that on our 'Profit Ratio Wrong' you get this option to choose the aggregation. We don't get this on our correct profit ratio because we set that aggregation within our calculation itself. The next question you are asking is: "Why are these same? These are the exact same number. You said they'd be wrong." But really, the reason they are acting the same is that we are at the row level of detail, the aggregation vs. non-aggregation does not matter here. However, if we go to Example 2 here, and recreate this view using our Continent Field in lieu of our Row ID, so we get a much higher of detail here, you'll see the problem illuminated brightly.

I'm going to bring profit out to 'Text' again, and double-click 'Sales,' and then I'm going to add my 'Profit Ratio Wrong' and my 'Profit Ratio Correct' tables as well. Now this is where you see the problem. So now instead of getting the correct response of 28.83% which I get using that aggregate calculation, my row level gives me crazy numbers that don't seem to make sense at the surface, but this is pretty easy to see what is going on here. If I bring 'Row ID' out here, again you'll see that the numbers are matching, so what we are doing here, if we keep a few of these rows just so we can demonstrate this, here we're getting the same numbers because we are at the Row ID.

If I were to highlight all of these, you'd see that I get, you know, the nine rows, they are matching, but what we are going to do here when I remove Row ID is these numbers will aggregate differently than these numbers. I will sum these up, getting that incorrect percentage, whereas in this case I take the sum of the profit for Africa and divided it by the sum of the sales instead of doing it for each row. So, here I get 113% because I'm summing up all of those distinct row level calculations whereas here I'm just taking 8,800 divided by 18,000 to get the correct number.

So just because the profit ratio has to be negative calculation in this context doesn't mean there aren't going to be contexts where row level calculations aren't useful. So in this case I want to see the average discounted unit price for my different categories and my data sets. So what I'm going to do here, since each of these rows might have different discounts, so I'm going to start by bringing out my categories. And additionally I'm going to bring out my Unit Price from my dimension. And I'll sort this appropriately. So we can see that my average unit price for copies and faxes is about \$750. But what is that when I take into account my discount? So we can see that my average unit price for copies and faxes is about \$750. But what is that when I take into account my discount?

So first of all, what does my discount look like? If I right-click this and choose 'Describe' you'll be able to see what this looks like. And it looks like it is a percentage so, between 0.0% and 25.0%, so I know that I can just multiply my year price by that percentage. So, I'm going to right-click my unit price here and choose 'Calculated Field' and since I just described my discount I know that it's just simply a straight percentage, I can just multiply it by that. So this is going to be unit price times that discount so Unit Price times a quarter at the most. What I want here is the actual discount price, so: Unit Price times 1 - Discount. And this is going to happen for every row because I know that instances of each product may have different discounts, so I'm going to call this line 'Average Discounted Unit Price'.

This is an appropriate calculation to do at the row level because I want to check this for each of my orders; this isn't something I want to do at the aggregate level because I don't care if one product has, you know, a 15% discount and another one has a 1% discount, I don't want to aggregate that, I want to do that at the row level and then aggregate those results. So the first thing I'll do here is I'll change this back to 'Average,' and now you can see, on average, I discount these by about \$30. So there are instances where row level calculations are going to be useful, so don't always just jump to the aggregation, though they are more common.

The final example we have here is talking about another type of calculation we have in Tableau. And these are known as 'Table Calculations.' There is going to be another video discussing these at a deeper level of detail, so please feel free to refer to that for more information as well. So in this case I'm going to start by using my order date, and I'm going to start by looking at my raw sales information, I'm going to bring out 'Order Date' and 'Sales.' You'll see we're using that Tableau hierarchy for the dates and building out to here automatically. But what I want to do is use my continuous date parts down here so I can see my sales by month over time. This is used for, I can tell visually in some cases, you know, where I went up, where I went down, but in cases like this it's somewhat difficult to tell whether I increased or decreased.

Now, how can I show this in a more visual manner? How can I see how much of a jump this was vs. to how much of a fall that was? I can do this with a quick table calculation, so if I hold down my Control key and click and drag 'Sales,' I'm just going to create a second copy of it up here on the row shelf, and on the second copy I'm going to left-click this dropdown, or similarly right-click the pill, and hover over 'Quick Table Calculation,' and I'm just going to choose

'Percent Difference.'

And you can see a couple of things right away, first of all the axis has changed to a percentage because of the calculation we have chosen, and second I didn't have to know the formula to create a percent difference. That was built in. What we're doing here is taking the information we receive back from the database, so all of this information here, which is known as the 'Table Data.' And running this table calculation on top of that Table Data. This explains why we get this one NULL because we have our first point here, we don't have any sales before January 2010, so there's no way to calculate that difference. But once we get to February, we know we have a large drop in sales, so we are able to calculate that difference. These calculations can be controlled in many ways, so I can change this to say: "Hey, I want to look at this by year."

The next date so I can start to go backwards; these are very flexible calculations and very powerful calculations. Additionally, they don't have to just be used on the rows shelf. I can hold down my Control key again and drag this and put it on 'Color,' so I visually really call out where the large differences are, both negative and positive.

So, this has been just a brief introduction in calculations in Tableau. If you have questions about specific function types, whether they be logical or number or table calculations, please refer to the next set of videos that discuss those in more detail as well as refer to the documentation and other knowledge-based articles we have on our website. Thanks for watching, happy analyzing!