On-Demand Training:
Aggregate Calculations
Transcript

Welcome to this video on aggregate calculations. You can download the Exercises workbook to follow along in your own copy of Tableau.

AGG Calculations

Have you ever seen an "AGG" at the beginning of a pill? AGG stands for 'aggregation' and tells us that the aggregation is built into the calculation rather than performed after the fact.

For example, we'll create a profit ratio:

- Double click in the Columns shelf, and we'll type

    - SUM(Profit)/SUM(Sales), and we'll hit enter

- Notice that the pill now begins with AGG instead of SUM.

- AGG means the aggregation is part of the calculation itself.

- For more information on how the Ratio Calculation is written, please refer to the video Aggregation, Granularity, and Ratio Calculations.

Attribute Aggregation

We know what AGG means, but what about ATTR? This is another common thing seen with aggregate calculations.

Let's say we want to aggregate separately for different dimensions values – for example, we want to average Corporate Sales but take the median of Sales for the other Customer Segments.

- Right-click in the data window and Create Calculated Field… We'll call this 'Segment Sales'

- And in the calculation editor, I'm going to go ahead and paste in the calculation. Feel free to pause the video and copy it down yourself.

    - IF [Segment] = "Corporate" THEN AVG(Sales) ELSE MEDIAN(Sales) END

    - This says, if the segment is corporate take the average of sales. Otherwise, take median sales.

- But Tableau tells us that there is an error in the calculation: cannot mix aggregate and non-aggregate comparisons. So we need to aggregate Segment, but how?

- By using the attribute aggregation – we'll wrap Segment in ATTR – make sure to include the parentheses

- And now the calculation is valid. We'll hit OK.

- Let's add Segment Sales to the view, and now we can see that the average corporate sales is much larger than the median sales for the other Segments.

But what did Attribute do?

- Attribute checks to see if there is only one value for a given field for all rows in result set. If there is only one value for the data selected, Attribute returns that value. If there is more than one value for that subset of data, it returns an asterisk.

    - Another way of thinking of the attribute function is as the equivalent to the logical test "if min equals

max then return that value"

We could also use MIN or MAX if we know there's only one value.   All of these aggregations are useful for aggregating dimensions to get around of errors like the one we saw.

Aggregated versus Record Level Calculations

Finally, what about aggregated versus record level calculations?

Let's look at the Average Sale per Order

- We'll start by right-click dragging Sales onto the columns shelf – or option-click dragging on a Mac. This allows us to choose an aggregation. Select Average and click OK

- This is the average sale per row of the database – which consists of product-level transactions, not order level. We want to group multiple items sold into a single sale for a given order and to take the average of that, not a record-level average.

- To do this, we'll create a calculated field.

  - Right-click, select Create a calculated field and we'll name it 'Average Order Sales.'

  - Averages are calculated by summing a set of values and then dividing by the number of values, so we'll do the sum of Sales divided by the count distinct of Order ID.

    - SUM(Sales)/COUNTD([Order ID])

  - Be sure to do a count distinct and not a simple count – doing a count would count the number of records rather than the number of distinct orders,

- Now we'll bring Average Order Sales to the columns shelf.

We see that these values are larger than the AVG(Sales) values, which makes sense because we are taking the average of values summed to the order level rather than the record level.

Conclusion

Thank you for watching this video on topics pertaining to aggregate calculations. We invite you to continue with the On-Demand Training videos to learn more about using Tableau.