On-Demand Training: Additional Data Blending Topics Transcript

Welcome to this video on Additional Data Blending Topics. You can download the Exercises workbook to follow along in your own copy of Tableau Desktop.

Data Blending and Calculations

Performing calculations across two different data sources can be slightly different from creating an ordinary calculated field.

Our current view shows a state-by-state comparison of sales data from Office City and Coffee Chain. We see sales information for every state from the primary source, Office City, in blue, but we are missing sales information for several states in Coffee Chain, in orange.

What if we want to view our combined sales for each state? In order to combine our sales from both data sources, we need to create a calculation.

- Verify that the Office City data source is selected,
- go to the Analysis menu,
- and select Create Calculated Field.
- Note the indicator tells us which data source we are creating this calculation on.
- Drag in Office City Sales
- And type a plus (+)

Next, switch to the Coffee Chain data source.

- Drag in Coffee Chain Sales.

Working with Aggregated Data in Calculations

Notice that the Coffee Chain Sales looks different from the Office City Sales in the formula box.

First, the Coffee Chain Sales also contains the name of the data source. This format is designed to help users distinguish between different data sources in a calculation, so you always know which data source a field is coming from.

Second, the Coffee Chain Sales is automatically wrapped in a SUM function. Data from the secondary source must always be aggregated, and SUM is simply the default aggregation.

As we can see, we have an error because the fields used in the calculation must either be all aggregated or not. To resolve this error, let's wrap our Sales from Office City in a SUM function. Our calculation is valid, so we'll rename this "Combined Sales" and click OK.

We do not see our calculation in the data window; this is because we're on a different data source than the one the calculation was created on. If we switch to Office City, we now see that calculated field.

Working with Null Values

Now we are ready to test our calculation.

- Drag the Combined Sales field onto columns.

- Hmmm. Notice that the Combined Sales data is not displayed for the states where Coffee Chain Sales is null.

When you create a calculation across two data sources, it is common to get some null results, because adding a null to a normal value yields a null value. However, we know that our nulls from Coffee Chain are actually zeros, so we need to adjust our calculation so it accounts for this.

- Right click on Combined Sales and Edit.
- We'll wrap the Coffee Chain Sales in a ZN function, this stands for "zero-null."

Now our Combined Sales column accurately displays the combined sales information for Office City and Coffee Chain.

Asterisks When Blending

Another common scenario with data blending is dealing with asterisks. When data blending, adding a dimension to the view sometimes results in an asterisk instead of the expected values. Why does this happen?

In our current view, we are using Office City as our primary data source, so it's color coded blue, and we're blending on State. Area Code is only present in the secondary data source. Because there is more than one area code for most of the states in this data, Tableau produces an asterisk to indicate the multiple values.

New Hampshire and New Mexico show values because there is only one area code in each state.

Asterisks Solution: Swap Primary and Secondary Data Sources

How can we address the issue of asterisks when data blending?

Try swapping the primary and secondary data sources in a new view.

We are now using Coffee Chain as our primary data source, and it's blue.

Verify that the Coffee Chain data source is selected and now we'll bring out Area Code to Rows. This is the expected behavior for Area Code, we now have an Area Code listed per state,

However, Office City does not have an Area Code field so Sales is aggregated to the level of the blend, State, and that value is repeated per area code.

If this is not an acceptable way to resolve the asterisks issue, consider joining your tables rather than blending and using table calculations or level of detail expressions to perform your analysis. For more information, please watch our on-demand training videos on these topics.

Conclusion

Thank you for watching this video on Additional Data Blending Topics. We invite you to continue with the On-Demand Training videos to learn more about using Tableau.