



Free Training Transcript:
Data Blending

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

What is Data Blending?

Imagine you own two retail chains, Office City and Coffee Chain. You capture your Office City data in one system and your Coffee Chain data in another. Data blending is one way of combining data from multiple data sources into a single view. Instead of joining the data at the row level like a cross-database join, data blending sends separate queries to the separate data sources and aggregates the results to a common level back in Tableau.

The Blend Relationships

Here we have two data sources in the data pane: Office City and Coffee Chain. We can easily switch between the two, and notice how the dimensions and measures change. Data blending requires at least one common field between both data sources, such as State. If the field names are different, but have common members, we can manually define their relationship.

For example, we know that Region from Office City and Territory from Coffee Chain both contain the values 'Central,' 'East,' 'South,' and 'West,' so let's define the two fields as equivalent. We'll go to the Data menu, select Edit Relationships, we can change the primary to be Office City, select Custom, and add a relationship: we'll chose Region and Territory.

Notice that Tableau has established a relationship between the two fields, and also lists the automatic relationship, state. Alternatively, we can rename the field in one data source to match the other. Let's slowly double click and rename Coffee Chain's Territory as Region. If they have the same name Tableau can create a relationship between both data sources automatically.

Now that our common fields are defined, we are ready to start blending.

How to Blend

First, select the Office City data source. We'll drag Office City Sales to the Columns shelf and we'll bring State to the Rows shelf. Note that there is now a blue check mark next to Office City in the data pane. Whenever we are connected to multiple data sources in Tableau, the first data source we bring out to the view becomes the primary, denoted by

this blue check mark. Now, let's switch to the second data source. And notice the orange linking icon next to State. Since we already brought State into the view, Tableau will automatically blend on it, denoted by the orange link.

There is also a gray, broken link next to the Region field. Because Region is a common field between both data sources, it is another potential linking field that is not being used in the current view. If we wanted to also blend on this field, clicking on the gray link activates it and creates the relationship. Let's complete our data blend by dragging Coffee Chain's Sales to the Columns shelf. We now see there's an orange check mark next to the Coffee Chain data source and on the pill in the view, this indicated that it's from our secondary data source.

How is this view built? Tableau is querying for Sum of Sales by State to each data source, then displaying those values together in the view. We see sales information for every state from Office City because Tableau returns information for all field members (states) in the primary data source, regardless of whether or not there is data from the secondary data source.

Swapping Primary and Secondary Data Sources

It's important to note that primary and secondary sources are determined on a worksheet-by-worksheet basis and are not maintained globally throughout the workbook. When we are on a new sheet, the data sources within the Data Pane do not have orange and blue check marks to indicate them as primary and secondary. The relationships we established in the previous worksheet are not carried over.

What happens when we swap our data sources and use Coffee Chain as the primary source and Office City as the secondary? We'll drag Coffee Chain Sales to the view first, then we'll bring out State. There are now only about half the states on the view compared with our previous example. This is because Coffee Chain has fewer states in its data set than Office City. Tableau is displaying all the states in Coffee Chain, and then it will pull in only the relevant information from Office City. This means that any state in Office City that is not in Coffee Chain will not be displayed in this view.

Conclusion

Thank you for watching this Data Blending video. We invite you to watch the video [Additional Data Blending Topics](#), as well as our other free Training videos to learn more about using Tableau.