



---

# On-Demand Training: Data Blending Transcript

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 Excel and your Coffee Chain data in an Access database.

How can we bring these disparate data sources together in order to perform an analysis across both retail chains?

Data blending allows us to combine data from multiple data sources into a single view in Tableau.

## Requirements for Data Blending

Before we dive any deeper, let's stop and discuss the basic requirements of data blending.

Take a look at the Data Pane in the upper left-hand corner. We have two data sources: 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. If you are familiar with database joins, these common fields function as a joining clause.

Instead of joining the data at the row level, however, data blending sends separate queries to the separate data sources and aggregates the results to a common level back in Tableau.

In our current example, we have a State field in both Office City and Coffee Chain. These fields have at least one common member.

If the field names are different, but have common members, we can manually define their relationship.

We know that Region from Office City and Market 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,
- let's change the primary to be Office City,
- select Custom,
- and add a relationship
- we'll chose Region and Market. .

Notice that Tableau has established a relationship between the two fields, and also lists the automatic relationship, state. We'll cancel this.

Alternatively, rename the field in one data source to match the other. If they have the same name, like State, Tableau will create a relationship between both data sources.

Let's right click and rename Coffee Chain's Market as Region.

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 onto the Columns shelf and State onto the Rows shelf.

- Note that there is now a blue check mark next to Office City in the data window.
- 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 on to Columns.

- We now see there's an orange check mark both next the Coffee Chain in the data window and on the pill
- This allows us to distinguish between our primary and secondary fields on the view.

## Data Blending: Behind the Scenes

What's going on in this view?

- Tableau uses the orange link field (State) to determine which rows of data are brought in from each data source.
  - By blending on State, we have told Tableau to check for shared members, or states, across both data sources.
- We see sales information for every state from Office City because Tableau returns information for all field members in the primary data source.
- Notice, however, that we are missing sales information for several states in the Coffee Chain data source.
  - When Tableau queries the secondary data source, it only returns information for field members that are shared with the primary.
  - As a result, there are blanks for states like Alabama and Arizona that are only present in the Office City primary data source.
- We can think of this as a simulated left join.

## 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 relationships we established in the previous worksheet are not carried over.
- This allows us to create different types of data blends on a worksheet-by-worksheet basis.

What happens when we swap our data source 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.
- Now we'll switch to Office City and bring out Office City Sales.

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 data for all the states in Coffee Chain, and then pulling in information from Office City only for the relevant states.

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. Please watch Additional Data Blending Topics, as well as our other On-Demand Training videos to learn more about using Tableau.