tableau®

Free Training Transcript: The Tableau Prep Interface

Welcome to this video on the Tableau Prep interface. You can download the data set and packaged flow file underneath the video. We're working with weekly data for bestselling books.

Tour of the Interface

Tableau Prep lets you clean, shape, and combine data to prepare it for analysis.

When we open Tableau Prep, we begin on the start screen. To the far right we have the Discover Pane, with educational resources and links to relevant content. In the middle, we have links to previous flow files and samples. To the left is the Connections Pane.

We can select a data source from the Connections Pane and drag a table out. This creates the first step in the flow, an Input step. Every flow must begin with an input step, though there can be multiple inputs in a flow. When we selected a step in the flow pane, its related pane will open below.

We can configure the input by choosing to bring in a single table or create a wildcard union. We can also configure the sampling, if our data source is large enough to merit being sampled. To the right we see the fields in the data and can pick exactly which fields to bring into Tableau Prep. We can also set filters here, limiting what data is brought into the rest of the flow. For more information on bringing data into Tableau Prep, check out the video on The Input Step.

To add another step, we click the plus on the previous step in the flow. We can choose from a cleaning step, an aggregate, pivot, join, union, or output.

A cleaning step appears in the flow as a simple bar. Below, we have the profile pane, which show us the fields in the data as cards. Each card displays the values in that field and the number of rows for each value, providing insight into the structure and relationships in your data. For more information on the profile pane, check out its video.

Cleaning operations are found in the menu for each card or the contextual toolbar. We can also do some cleaning by direct interaction, such as changing the data type or renaming the field. Each type of cleaning operation we perform will appear as an annotation above the cleaning step in the flow pane and will be tracked in the Changes to the left. For more information on cleaning in Tableau Prep, check out the video on The Cleaning Step.

Below the profile pane is the data grid, which shows rows of the data, similar to a spreadsheet view.

Let's open a more complicated flow and go through some of the other steps we can do in Tableau Prep.

Types of Steps

Here, we have a complex flow. We can break it down into parts--there are a series of pivots that restructure data for January's Bestsellers, bringing it all back together with a series of joins. There's a union that combines several weeks of data for February. March is also pivoted, and all three months are brought back together with a union. Another data source brings in sales information and aggregates it from transactional data to book-level data, then joins it with the bestsellers. Finally, movie adaptation data is also brought in via another join, and an output is generated for the fully combined data source.

Each type of step in a Tableau Prep flow is present here. We can click any part of the flow to bring up the panes associated with that step.

First, let's look at a Pivot, indicated by this "columns to rows" icon in the flow. Fields are brought from the left side list of fields to the Pivot Values drop area. The results, that is, the pivoted fields, are shown on the right, with both a profile pane view and a data grid view. For more information on how to perform a pivot in Tableau Prep, check out the video on The Pivot Step.

Next, let's click to an Aggregation step, indicated by a Sigma icon in the flow. Fields are again displayed on the left, and brought to either the Grouped fields area or the Aggregated Fields area, depending on how each field should be treated by the aggregation. For more information on how to perform an aggregation in Tableau Prep, check out the video on The Aggregate Step.

Combining data can be done by unioning (adding more rows with the same column structure) or joining (adding new columns to existing rows of data). Let's look at a union step.

A union shows the inputs going into the union on the left, and on the right, we see the profile pane and data grid view of the results. Each field's card displays a colored bar, showing which data inputs are present in that field. If we need to merge fields, we can do so with a simple drag and drop. There's also a "Table Names" column generated by Tableau that contains information about which data source each row came from. For more information about how to perform a union in Tableau Prep, check out the video on The Union Step.

Joins are another way of combining data. A join step provides a visual interface for configuring the join. To the far left, we can build the join clause or clauses, choosing which fields to join and how. We can select the Join Type by clicking on the diagram. Under the join type, we see the summary of results, showing us what is being matched (included in the results) or mismatched (excluded from the results). If we change the Join Type, we see the summary of results update.

Just to the right of the Join Configuration Pane, we see a breakdown of the Join clauses. Values in red are unmatched, and values in black are matched. Further to the right, we see the profile pane and data grid view of the results. For more information about how to perform a join in Tableau prep, check out the video on The Join Step.

Finally, once the data has been cleaned, shaped, and combined to our satisfaction, we must output the data. Saving the flow file (.tfl) will save the flow itself, not generate an output of data. A packaged flow file (.tflx) includes the flow and extracts of flat files (though not data from database connections), and is not the same as outputting the finished, cleaned data. To create the finished data source, an output step must be added to the flow and the flow must be run.

In the output pane, the right side shows the data grid preview of what the finished data will look like. To the left, the output can be configured. It can either be saved to a file or published as a data source. The file can be named and set as either a .hyper extract, a .tde extract, or a .csv, and we can specify where to save the file.

Clicking Run Flow will execute the flow and generate the output file or publish the data source. Note that if a flow has multiple outputs, they can be run independently or simultaneously. For more information, check out the video on the Output Step.

Conclusion

Thank you for watching this video on the interface of Tableau Prep. We invite you to continue with the free training videos to learn more about using Tableau.