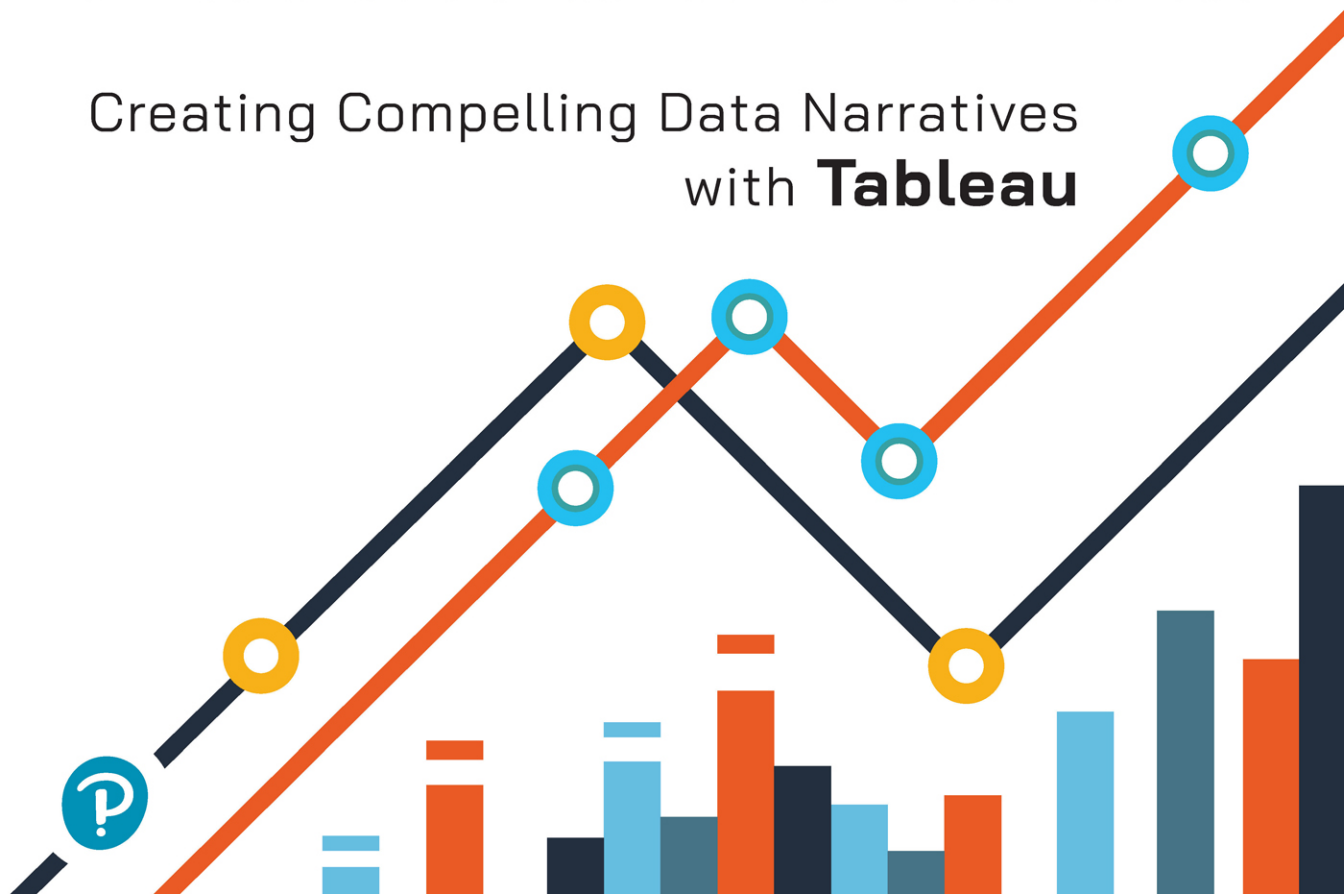
LINDY RYAN

VISUAL ANALYTICS FUNDAMENTALS

Creating Compelling Data Narratives with **Tableau**



Visual Analytics Fundamentals

From here, as you begin to explore and analyze your data, you can selectively replace data sources at the worksheet level (in the past, replacing data sources would apply to all worksheets in the Tableau workbook). To replace the data set used, return to the Data menu and select Data > Replace Data Source. This option enables you to replace the data source with any other connected data source, at either the worksheet or workbook level (Figure 3.13).

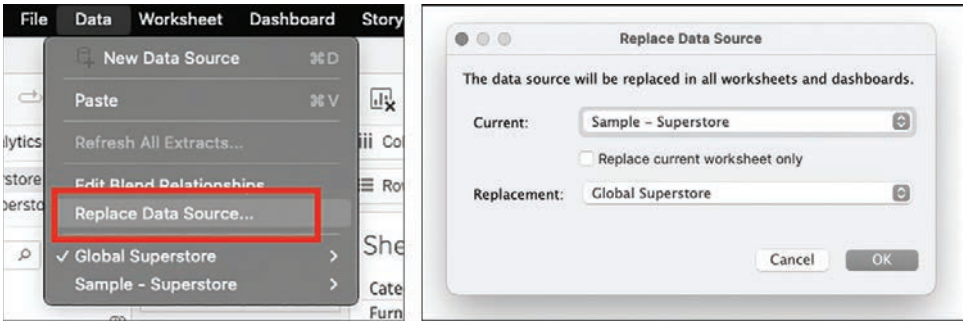


Figure 3.13 Data sources can be changed at either the workbook or worksheet level.

Basic Data Prep with Data Interpreter

Tableau Desktop delivers some features to help automatically reshape files to get them ready for analysis in Tableau without relying on additional tools (such as Tableau Prep) or other data prep tools. One is Data Interpreter, Tableau's built-in tool for preparing data for analysis. While many of the Excel files we'll use in this text are already nicely formatted and ready for analysis in Tableau, real-world data files are not always so well prepared. If you need to connect to a data table where minimal amounts of data prep is necessary, Data Interpreter is a handy resource.

When you connect to an Excel sheet in Tableau, the software can recognize issues such as missing column names, null values, and so on. To remedy these problems and clean the file for use in analysis, Tableau will suggest using Data Interpreter. (Refer to Figure 3.4 to locate the Data Interpreter option on the left pane of Data Source screen, directly between the list of data connections and the resultant tables.)

We will take a closer look at Data Interpreter in a later chapter, but for now it's sufficient to know that to use Data Interpreter, you simply click the check box to activate this tool. This executes a query to the Excel file and confirms its automated prep tasks, with a revised data preview pane addressing the issues Data Interpreter has identified. To get more specifics on what Data Interpreter has adjusted in the file, including a before-and-after view and an explanation table, click the link that is provided

following Data Interpreter's action to "Review the results." This opens an Excel file describing the changes. You can also clear the check box to undo these changes and revert to your original sheet.

After verifying the changes made to your data, you can go to your worksheet and begin exploring the Tableau interface and your data. You are ready to begin your analysis!

Navigating the Tableau Interface

Now that you have connected to some data in Tableau, you can click the prompt to Go To Worksheet and start getting to know the Tableau user interface (UI) in a more meaningful way. Like the Data Source page, the Tableau UI is a drag-and-drop interface that fosters rich interactivity between sheets, dashboards, and stories, allowing for in-depth visual exploration and powerful visual communication. Tableau is similar to Excel in that its files are called workbooks and the sheets inside the workbook are called sheets. Every Tableau workbook contains three elements:

- **Sheets:** For creating individual visualizations. Each workbook can contain multiple sheets—one for each data visualization you create.
- **Dashboards:** For combining multiple sheets as well as other objects like images, text, and web pages, and adding interactions between them like filtering and highlighting. Dashboards are great for looking at the interactions between multiple visualizations in a single view.
- **Stories:** These frameworks can be based on visualizations or dashboards, or they can be based on different views and explorations of a single visualization, seen at different stages, with different marks filtered and annotations added. However, stories are best suited to narrate the story in your data.

We'll cover dashboards and stories, and the differences between them, in more depth in later chapters. For now, let's focus on sheets and take a high-level view of the various areas of the Tableau worksheet canvas. As we begin to work directly with data to perform visual analytics and build visualizations, dashboards, and stories throughout this book, we'll explore these areas—and more—in detail through hands-on exercises. For now, this high-level overview is intended to orient you to the various aspects of the user interface.

As shown in Figure 3.14, the Tableau interface includes five basic elements:

1. Menus and toolbar
2. Data pane
3. Shelves and cards

4. The canvas workspace
5. Legends

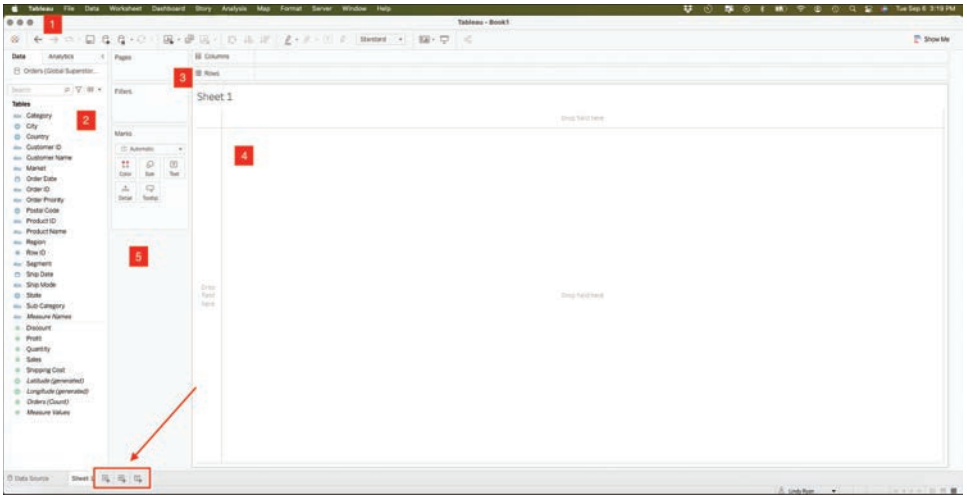


Figure 3.14 The Tableau user interface, a blank canvas.

Menus and Toolbar

Even though most of your Tableau work can be accomplished by interacting directly with the drag-and-drop canvas, the menu bar that launches with the software includes various menus that provide access to additional features and settings. This includes the File and Data menus, as well as the Worksheet, Dashboard, and Story menus, each of which contains specific controls for those canvas types. Additional menus, such as the Analysis, Map, Format, Server, Windows, and Help menus, contain even more functionality and controls.

At the top of the Tableau sheet is the toolbar, which is similar in concept to the ribbon in Microsoft Office products. The toolbar contains many powerful buttons that give you control over your Tableau experience and enable you to navigate from the data source all the way to story presentation mode. A few items of special note are highlighted here:

- **Logo:** The Tableau logo button brings you back to the original Connect to Data screen (clicking the icon from this screen returns you to your sheet).
- **Undo:** There is no limit to how much you can undo in Tableau, which is an important feature for exploration and discovery. The icon is grayed out until there is an action to undo.

- **Save:** There is *no automatic save* in Tableau. Be sure to save your work incrementally.

Another menu appears along the bottom of the sheet. This menu, similar in concept to a Tableau workbook, enables you to return to the Data Source screen; create new sheets, dashboards, or stories; and do things like rename, rearrange, duplicate, delete various sheets, and so on.

Data Pane

The pane on the left of the sheet is called the Data pane. It has two tabs: a Data tab and an Analytics tab.

Data

At the top of the Data tab is a list of all open data connections and the fields from that data source categorized as either dimensions or measures (discussed shortly).

Analytics

The Analytics tab enables you to bring out pieces of your analysis—summaries, models, and more—as drag-and-drop elements. We will review these functions later.

Shelves and Cards

Shelves and cards are some of the most dynamic and useful features of the Tableau UI.

- **Columns and Rows shelves:** Control grouping headers (dimensions) and axes (measures).
- **Pages shelf:** Lets you break a view into a series of pages so you can better analyze how a specific field affects the rest of the data.
- **Filters shelf:** Filters visualizations by dimensions or measures.
- **Marks card:** Controls the visual characteristics of a visualization, including encoding of color, size, labels, tooltip text, and shape.
- **"Show Me" card** (shown closed): A collapsible card that shows application visualization types for a selected measure and dimension.

Legends

Legends will be created and automatically appear when you place a field on the Color, Size, or Shape card. To change the order (or appearance) of fields in a visualization, drag them around in the legend. Hide legends by clicking on the menu and selecting Hide Card. Likewise, bring them back by selecting the Legend option on the appropriate space in the Marks card or by using the Analysis menu.

Understanding Dimensions and Measures

When you bring a data source into Tableau, the software automatically classifies each field as either a dimension or a measure. The differences between these two are important, though they can be tricky to understand for those who are new at analytics. Perhaps the best way to differentiate these two classifications is to think about them this way: Dimensions are categories, whereas measures are fields you can do math with.

Dimensions

Dimensions are things that you can use to group data by or drill down by. They are usually, but not always, categories (e.g., City, Product Name, or Color), and they can be logically grouped into strings, dates, or geographic fields. Dimensions can also be organized into Tableau groups and hierarchies, which we'll discuss shortly.

Measures

Measures are generally numerical data on which you want to perform calculations—summing, averaging, and so on. The definition of a field as a measure or dimension can be adjusted in the Data Source screen by clicking the data type icon. You can also change this directly in the sheet by either dragging and dropping a dimension to measure, and vice versa, or by clicking the drop-down menu by any field and selecting the Convert to Measure (or Dimension) option.

Continuous and Discrete

Generally, dimensions are discrete, whereas measures are continuous. We could break this down a little more into four types or levels of measurement:

- **Nominal measures** are discrete and categorical (e.g., for/against, true/false, yes/no).
- **Ordinal measures** have order but there are not distinct, equal values (e.g., rankings).
- **Interval measures** have order and distinct, equal values—or at least we assume they are equal (e.g., Likert scales).
- **Ratio measures** have order, distinct/equal values, and a true zero point (e.g., length, weight).

In Tableau, continuous fields produce axes, whereas discrete fields create headers. Continuous means "forming an unbroken whole, without interruption." Discrete means "individually separate and distinct." Be sure you understand the difference between these mathematical terms. Text and categories (dimensions) are inherently discrete. Numbers can be discrete if they can take only one of a limited set of distinct, separate values (e.g., a rating). Numbers, including dates, can be continuous if they can take on any value in a range.

COLORFUL PILLS

When a field is brought from the Data pane and dropped into the Rows and Columns shelves, Tableau creates a "pill." These pills are color coded: Blue pills represent discrete variables, whereas green pills are continuous. The data type icons also reflect these color codes (Figure 3.15).

Figure 3.15 Color-coded pills reflect continuous (green) measures and discrete (blue) dimensions.

Summary

This chapter introduced the Tableau product ecosystem. It then took a high-level view of the Tableau user interface, including connecting and preparing data and the core functionality of the Sheets canvas. In future chapters, you will put this knowledge into practice as you begin working hands-on with this functionality.

The next chapter addresses the importance of context in visual analytics.