Chapters to Go
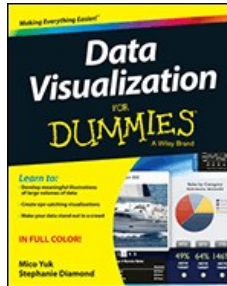
Data Visualization for Dummies

by Mico Yuk and Stephanie Diamond
John Wiley & Sons (US). (c) 2014. Copying Prohibited.

Skillsoft

Chapter 5: Adding a Little Context

In This Chapter

- Understanding the role of text in data visualizations

- Getting the most out of text

- Analyzing text

Remember the saying "A picture is worth a thousand words"? Well, although this saying may hold true for museum art, it doesn't apply to most data visualizations. One of the main reasons why users request data visualizations is to get away from tons of reports with loads of data and text that either take too long to interpret or are too difficult to understand.

If you want to create truly powerful data visualizations, adding some context in the form of text is one of the most effective ways to communicate your data. Take a stop sign, for example. Have you ever noticed that besides the unique shape and color of it, the text STOP is right in the middle? (See Figure 5-1.) Globally this is probably one of the most recognized signs — identifiable just by its shape and color. The text on the sign text just makes it that much easier to recognize what the sign is telling you. There is no room for misinterpretation; the stop sign is a good example of using text to add smart context to a visualization.

You should apply the same practice to your data visualizations. You shouldn't make the assumption that a visualization that's clear to one user will be as clear to 100 other users. This is where adding some context to your data visualization becomes important. Adding context in the form of text is a foolproof way to ensure that all your users get the same message.
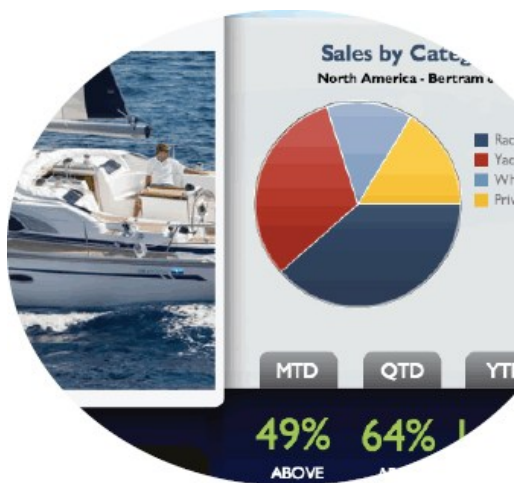

Figure 5-1: Users recognize what a stop sign is communicating based on the shape, the color, and the text.

Adding any form of text to data visualizations may seem to be counterintuitive, but this chapter shows you that with the right application, text can be a powerful addition to any data visualization. The chapter also explores the notion of text analysis, because applying it to your data visualization may make the difference between displaying intelligent data and displaying useless data.

Making Text Useful

One of the easiest ways to add context to any data visualization is to add text on or around a given visualization. Although you have an infinite number of ways to do this, this chapter focuses on a few specific techniques that make the process simple.

Before you go crazy adding text to your visuals, however, you should adhere to certain rules to avoid cluttering your visuals:

- **Use as few words as possible**. Text added to any data visualization must be complementary. Space is usually very limited, so your goal is to explain your visual with as few words as possible. E-mail marketers do this all the time, sometimes spending hours or even days on a single subject line, hoping to gain the highest e-mail open rate possible. The open rate is calculated by expressing the number of e-mails opened as a percentage of total e-mails.

  This process usually requires careful word choices, consideration of the limited space available in a single line, and multiple split tests to confirm the effectiveness of the chosen subject line. Done correctly, the process can increase e-mail open rates astronomically. The same is true of adding a line or two of text to a given data visualization. Done correctly, it can virtually eliminate misinterpretation of the data.

  Technical Stuff Split testing is a marketing method that splits e-mail subscribers into two groups. Each group is sent a separate e-mail that is then tracked and compared based on specific metrics. The one with the better metrics is used.

- **Stick to simple words or single characters**. Using simple words is the key to providing your users an at-a-glance understanding of your visual and accompanying text. Say that a number has "gone up" or use an up arrow (↑) instead of saying that something has "increased," or say that something is "bad" rather than "negative." These examples present information in a way that's easy for anyone to digest.

  Remember Simplicity is key when it comes to choosing words.

  Tip Avoid using acronyms that require some existing knowledge or additional action to interpret.

- **Use single lines of text**. Besides keeping text simple, you need to keep it short. Just the sight of multiple lines of text may make a user hesitant to start reading that text. Your goal should be to use a single line of text that's easy to read and interpret.

- **Apply color sparingly**. Although adding RAG (red, amber, and green) alert colors can be useful, adding random colors to make text stand out against your visuals isn't a good idea. In general, data visualizations are color-rich, and adding text in a color that contrasts with the visual causes the text to compete with the visual. When you're a beginner, try to stick to text colors such as black, blue, and gray. Read more about using color in Chapter 10.

  Tip Colors are used as alert colors in data visualizations. Think of a traffic light: Red means "Stop, something is wrong!"; yellow means "Proceed, but with caution!"; and green means "Go, everything is OK!."

- **Know your data**. One cardinal rule about adding any form of text to describe a visualization is to make sure that the text applies to *every* scenario of the data being displayed — especially if the data is dynamic. This means that you can't include a static set of text for only one scenario of the data visualization, because the text won't be applicable when the data changes along with the scenario. This situation is where dynamic text comes in handy. You can read more about dynamic text in "Exploring Text Analysis," later in this chapter.

Now that you know the five main rules for adding text to data visualizations, it's time to explore a few techniques for applying those rules effectively.

Adding text labeling

The quickest way to apply text to any data visualization is to add labels. Labels describe what the user is seeing. They're easy to use, easy to read, and easy to apply.

Table 5-1 lists three types of text labels that you can add to any data visualization.

Table 5-1: Text Label Types

Label Type	Description
Title	Usually placed at the top of, inside, or alongside a specific visual. The title describes what the visual is displaying — Economic Growth in 2012, for example. Some higher-level titles may describe a section that includes multiple visuals.
Description	Normally located alongside or below a specific visual. These labels use a single word, text fragment, or sentence to describe the behavior of the visual. A good description sums up the message of any data scenario in a single sentence.
Value	May include a single number, letter, or combination of both in or on a given visual. The label usually stands alone or refers to a specific part of the visualization. A single visualization may have multiple values.

Tip Most good data visualizations include all three types of labels.

Remember Labels should be strategically positioned (see the next section), sized, and worded so that the text complements the visuals without overshadowing them. Going overboard with text can lead to something your users will see as just another report with some visuals.

Considering text positioning

You may have heard the saying "It's all about where you place it," which could not be truer when it comes to adding text to data visualizations. Wrongly placed text can lead to misinterpretation of the data.

You should follow two best practices in placing text on your data visualizations:

- **Keep all text horizontal**. Nonhorizontal text may confuse some of your potential users. We strongly advise that you stay away from positioning your text vertically or diagonally. Never force a user to tilt "his" or "her" head to read something! Stick to the basics by keeping all text horizontal.

  Remember Keep in mind that text is included to complement the visual, not overshadow or complicate it.

- **Keep all text within range of the visual**. The human eye naturally associates any text within a few centimeters of it as being related to it. In general, ensure that all text is within the range of the visual it's related to, because positioning text too far from the targeted visual can lead to confusion.

  Figure 5-2 shows two basic bar charts. Both charts have a title label (refer to Table 5-1, earlier in this chapter), which is one of the most intuitive and common uses of text in any data visualization. The chart on the left has an easy-to-read horizontal label; the chart on the right has the same label tilted. Do you have to tilt your head to read the chart on the right? That label might be mistaken for an axis label.
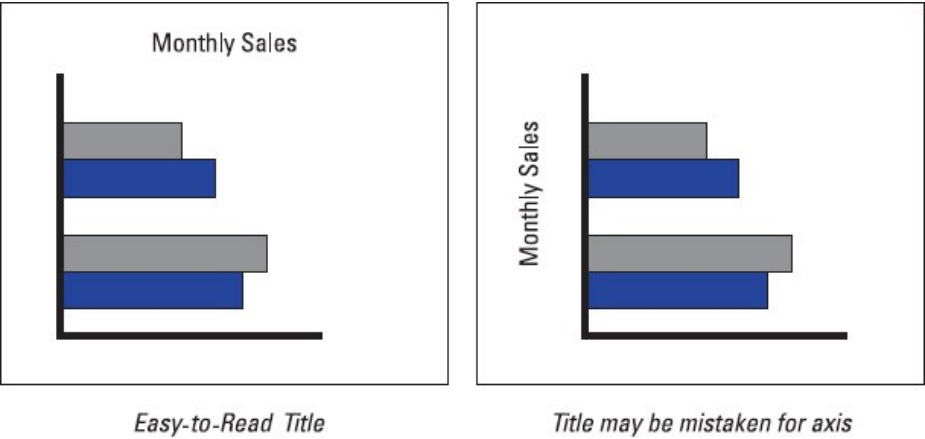
Figure 5-2: A title is in two different places on these charts.

Choosing text fonts

Choosing the perfect text font for any given data visualization can be a tricky process. Although increasing the size of a word may convey emphasis or importance, it's not always the best route as it may overshadow the visuals.

Font sizes and formats

Here are some ideas to consider when you choose font sizes and formats:

- **Make important text a little bigger**. A general rule of thumb is that low-hierarchy text should be in a smaller font than high-hierarchy text. Figure 5-3 shows how the font sizes of various labels in a hierarchy may vary based on level.

Figure 5-3: Fonts vary in size based on level.

- **Consistency is key**. If you decide that all your main titles will be in a size 30 font, for example, you must ensure that all your main titles are exactly the same size.

Remember The key factor is to have consistency for each individual label type. Consistent label sizes for each level are easy for users to grasp, and the sizes help users identify certain areas of your data visualization at a glance.

- **Avoid using all caps**. The era of text messages and social media has created a new set of rules for using text. Today, using all caps indicates that someone is shouting, which usually evokes hostility and defensiveness. All caps are also harder to read, as all the letters are the same size. In addition, readers may gloss over or skip all-caps words.

  It may be fine to use a simple all-caps title such as SALES EXECUTIVE DASHBOARD, but don't use all caps in labels throughout your data visualization. In general, you have very little to gain from using all caps, so it's best to avoid the practice.

Font types

Thanks to web browsers such as Internet Explorer and Google Chrome, most people have become accustomed to a few standard fonts, such as Verdana, Garamond, Arial, Helvetica, and Times New Roman.

These fonts are somewhat boring, but we highly recommend sticking to one or more of them, for two reasons:

- **Each user's web browser may be different**. Data visualizations are commonly hosted on company sites or portals that must be viewed in a web browser (unless, of course, you're developing a static image like an infographic). No matter what fonts you use to create a data viz, each user's web browser automatically defaults displaying any text using one or more of the browser-safe fonts mentioned earlier.

  Warning! If you use unusual fonts, your visualization may look distorted or become unreadable when your fonts are replaced with browser-safe fonts. It's safer to design your visualization with the common fonts.

  Tip One way to get around having to use browser-safe fonts is by posting your visualization in an Adobe PDF or Microsoft PowerPoint format. These file types are not affected by the browser, but these options are not available in all tools.

- **Avoid fancy or custom fonts**. Using fancy or custom fonts may make it harder for someone else to understand or edit your data viz, especially if you use nonstandard or specialized fonts that need to be individually downloaded and installed on a desktop. Yes, custom fonts may look great, but they are simply not worth the risk of having the user not being able to view them.

Tip We recommend using Verdana and Garamond, both of which are on the extended list of browser-safe fonts. They add a little bit of flair without having the straight military feel of Arial.

Choosing text color

As we mentioned earlier in this chapter, you should choose a font color that's easy to read, such as black, gray, or dark blue. You should use this color scheme consistently throughout your data visualizations. The last thing you want is a bunch of color-rich visuals that are accompanied by colorful text; users would have a hard time deciphering what's important!

Consider a few guidelines when choosing text colors:

- **Consistency is king**. This won't be the last time we'll tell you to be consistent. Maintaining consistency of font color for specific label types is an absolute must. If you choose to put all descriptions in a dark gray color, then it's important that all descriptions receive this same treatment. Newbies often get the urge to change the color of particular text in an effort to bring more attention to it. They don't take into consideration how the change affects the entire data visualization. Avoid falling into the trap of using too many colors.

Avoid data overload

Mainstream advertisers often use the term *data overload to* scare the user community. Just type **data overload** in a search engine, and you'll see a series of data visualization software choices that are guaranteed to help even the most novice of users avoid the dreaded data overload syndrome by supposedly automating the generation of "useful" data visualizations! This promise is often an advertising ploy, though, so it's very important not to fall for this marketing trap.

A good rule of thumb is to ensure that your text-to-visuals ratio is 1:5, which means that you limit the use of text to roughly 20 percent of your overall data visualization. This ratio is a surefire way to ensure that you add value rather than clutter to your data viz.

- **RAG colors are sacred**. The use of red, amber, and green in all data visualizations is sacred, especially in the financial world. When it comes to data visualizations, these colors are automatically interpreted as meaning that something has gone wrong (red), something is about to go wrong (yellow), and everything is fine (green). These alert colors are used in data visualizations to provide notification when a measurement is heading to a good or bad status. When it comes to text, you should use only RAG colors to indicate alerts.

- **Bright colors are hard to read**. What happens when you see a neon-colored sign? It definitely captures your attention at first. But imagine having to look at the same sign every day! After a short period, you find that the sign isn't as easy to look at as it was when it first caught your attention. For this reason, avoid using bright-colored text in your data visualizations. Over time, muted colors will prove to be much easier for your users to read.

  Tip In some situations, however, a specialized color is required. See Chapter 10 for more information about those situations, which occur at a later stage in your data visualization journey.

Exploring Text Analysis

One of the most powerful ways to add context to your data visualizations is to include text analysis. In the Business Intelligence Dashboard Formula world (BIDF; see Chapter 1), text analysis is defined as a short sentence or fragment that contains a combination of static text and dynamic values that describe one or more visualizations.

Figure 5-4 shows a simple text analysis. The underlined words are *dynamic*, which means that they change based on the behavior of the data.

Sample Text Analysis Statement

ALERT: Expenses <u>have **exceeded**</u> monthly budget by <u>**37%**</u>!

Figure 5-4: Chart with text analysis statement with both static and dynamic text (underlined) that changes with the data.

Think of text analysis as being a real-time description of what's happening with the data. As the data changes, so does the descriptor. The trick is to create a statement that's smart enough to suit all scenarios of the data as it changes. Let's take a closer look at Figure 5-4 and how we built it.

First, you must know the data, as we mentioned earlier when we suggested rules you should follow for producing uncluttered visuals. For instance, in the example in Figure 5-4, we know that the chart is displaying expenses on a monthly basis. Therefore we establish the static parts of our sentence:

(ALERT): Expenses (dynamic text #1) monthly budget by (dynamic text #2).

Next you establish all the possible scenarios of the dynamic text (marked as blanks in the preceding example). In the case of Figure 5-4, we know that the possibilities for dynamic text #1 are that the expenses can exceed, not change, or remain below the monthly budget. Hence this sentence could have three different scenarios:

- **Scenario #1:** ALERT: Expenses (have exceeded) monthly budget by (dynamic text #2).

- **Scenario #2:** ALERT: Expenses (have not changed) monthly budget by (dynamic text #2).

- **Scenario #3:** ALERT: Expenses (remain below) monthly budget by (dynamic text #2).

Lastly, we need to fill in dynamic text #2. To calculate that rate for scenarios 1 and 3, we do a simple formula to find out by what percentage the expenses exceed or are below the budget.

Be sure to note the importance of not only covering all scenarios of the data but also changing the wording as needed to ensure the sentence makes sense. This is obviously a simple example, but it's a great place to start!

Determining what makes text analysis so important

When they first view a data visualization, most users do a two-part analysis. At first glance, they scan the visuals to get a sense of how the measurement is performing. Then, if time permits, they take a closer look at each visual to try to interpret what they're seeing. Their brains do a series of scans, capturing numbers, words, and colors to gain a clear idea of the message that the data is conveying.

You can use text analysis to eliminate this two-part process, however, by providing users with everything that they need to know in a single glance.

Text analysis serves two important purposes:

- **It prevents misinterpretation**. If 50 people see a single visualization, chances are that none of them will interpret the visualization exactly the same way. By providing text analysis, you essentially help users interpret the visualization clearly and make informed decisions.

- **It saves time**. Providing a clear, text-based interpretation of what a visual portrays saves users time, because they no longer have to stare closely at the visualization to gain a deeper understanding of the data.

Building a text analysis statement

How do you build a text analysis statement? Over the years, we've broken down thousands of text analysis statements and discovered that most of them contain the same four elements:

- **What:** This element refers to what you're actually measuring and usually is some form of key performance indicator or metric. Sales, expenses, margin, performance, and retention are all examples of measurements that you include at the beginning of your statement to let the user know what you're referring to.

- **What happened:** This element describes the behavior of the metric and uses measurement terms such as *high, low, up, down, increasing*, or *decreasing*. This behavior can also be represented with symbols, such as arrows.

- **By how much:** This element describes the magnitude of change in some numerical value.

  Remember Gauging the magnitude of any change is one of the best use cases for text analysis.

- **When:** A good guideline (when possible) is to include some reference to a time period to enable the user to gauge the effect of the change.
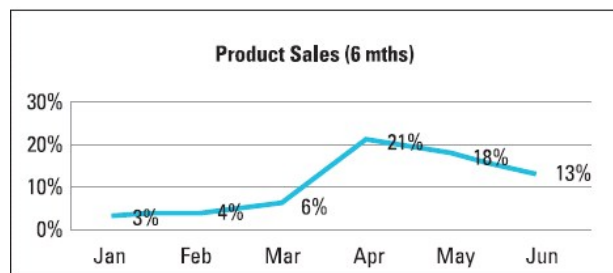
Figure 5-5 shows two simple text analysis statements that describe the same chart. Notice that the underlined dynamic words change as the data in the chart changes.

Positive Text Analysis
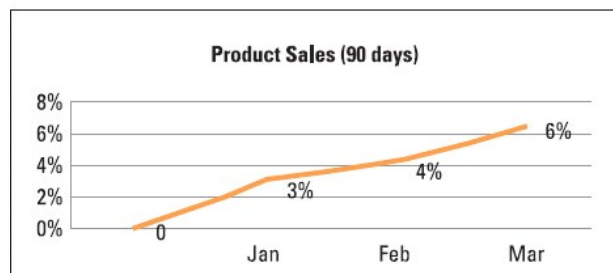
Sales went **up** by **6%** over the last **3 months**.

Negative Text Analysis

Sales went **down** to **13%** in **June**,

Figure 5-5: Examples of positive and negative opinions.

Incorporating dynamic text analysis into your data visualization makes your data truly intelligent. Including dynamic text analysis is low-hanging fruit with regard to your return on investment. It's easy to do and will delight your users. Some users may even think that you have magical data visualization powers!