Five Common Data Visualization Mistakes and How to Avoid Them

Giant leaps in science and technology over the decades have had a major impact. Vinyl records, grainy movies, and low resolution video have long been consigned to history, pushed out by HD content on tiny handheld devices wherever and whenever we want it. More and more we expect everything to be bright, colorful, and above all, instant.

So it is with information technology with high resolution monitors, projectors, and handheld displays making it far easier to access and consume the data needed in our working lives. More than that, Business Intelligence applications, such as TIBCO Spotfire®, are becoming increasingly adept at displaying data in new, visually compelling and immediate ways, a triumph over plain text and tables.

Software that analyzes and visualizes data is fast becoming the norm. These tools produce great visuals on demand and make it far easier to spot potential problems and capitalize on trends that might otherwise be missed. But, great visuals aren't always what you end up with. Human error, impossible deadlines, and a lack of understanding about what makes a great visual can stand in the way.

You can't always prevent human error or impossible deadlines, but with the aid of this ebook, you can add to your data visualization knowledge by understanding five of the most common mistakes and how to avoid them.

- **Visualizations should add value.** If they don't, leave them out.

- **Associate size with importance.** Use separate, possibly smaller visuals to illustrate less important data. Use separate charts to answer potential follow-up questions rather than trying to cram everything into one.

- **Use appropriate language.** Avoid technical jargon if it's not familiar to your audience and use scales they can relate to: hundreds, thousands, or millions to represent large numbers, and local currencies when charting financial data.

- **Use text when it helps.** For most audiences, a written explanation will convey the information faster than a visual alone. At the other extreme, a visual that requires a lengthy explanation will need to be made simpler.

- **Use the Spotfire Recommendations wizard to see best-practice visualization rules in action.** Recommendations will suggest visualizations that match the data, which can save a lot of time and effort.

- **Use the flowchart below.** When visualizing more complex information, it may help.

- **Use color sparingly**. Use color to highlight important results and categorize values, but not to represent quantities. Coloring areas green where yes votes came in strong and red where the vote was decidedly no is OK, and you can probably get away with using darker shades to show larger values; But, expecting users to remember which color goes with which percentage is asking too much.

ERROR 1: VISUALIZING JUST BECAUSE YOU CAN

One of the most common visualization mistakes is the assumption that any visual is better than none, or that consumers of a chart or dashboard will immediately understand what it is you're trying to say, no matter how obscure the data or the way it's presented.

What you want to convey through a visualization may be obvious to you, but don't assume it is to others.

AVOIDING THE MISTAKE

Think about what it is you need to convey and make sure the data supports it. (Don't start with a goal and try to manipulate the data to prove it.) Next, think carefully about your audience, and their ability to understand the information you want to get across. Think about the questions they are most likely to ask and want answered by the visualization.

When you know what you want to achieve:

- *Filter and highlight data to make a point.* If you leave conclusions up users to work out, chances are they won't bother or will get it wrong.

- *Get straight to the point.* Put the most important chart first, and when building a dashboard, put it at the top and make it big.

- *Try to tell a story with your visuals.* Position story elements from left to right and top to bottom.

ERROR 2: USING THE WRONG CHARTS

There may be statistical whizz-kids in your audience, but this cannot be assumed. As a species, humans are notoriously bad at understanding what numbers represent. Graphs can really help, especially when comparing big numbers, sequences, and large collections of data, but you still need to make sure you use the right kind of chart. Use one that takes a different approach from how users imagine data, and you make life unnecessarily hard.

AVOIDING THE MISTAKE

As a general rule, stick with basic line, bar, pie, and scatter charts and keep them as simple and clean as possible. Avoid the temptation to use 3D which may have a greater initial impact but can both distort the data and make it a lot more difficult to interpret.

Basic rules to follow:

- To highlight and compare values, use a bar chart.

- To show change over time or how one value affects another, use a line graph.

- To show how a whole is divided up, use a pie chart, but only for up to seven categories. For more than that, consider a tree map or a bar chart with 0-100% axis.

- To show relationships between values, use a scatter chart.

What are you trying to display?
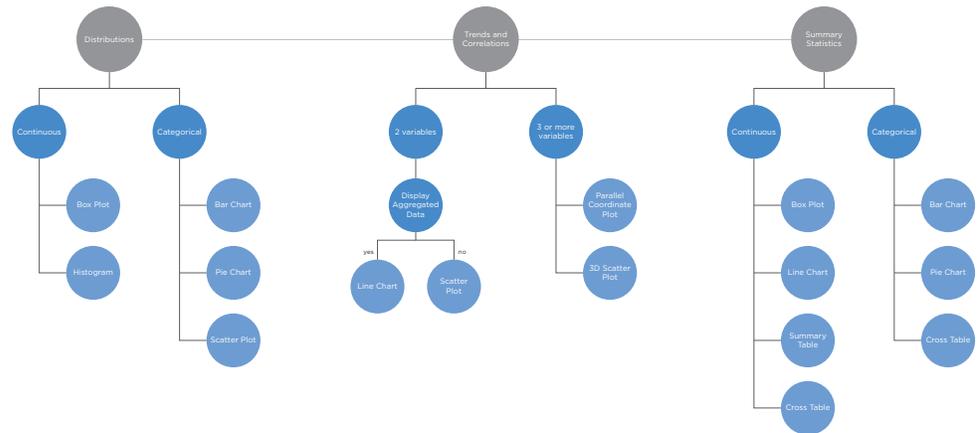
ERROR 3: TOO MUCH, TOO SOON

Large and complicated visualizations can be hard to comprehend. Representing large numbers is not easy when differences between numbers are small. For example, in a sales chart, a difference of less than 1% between regions could be significant, but that difference would get lost in a graph with a scale measured in millions of units.

AVOIDING THE MISTAKE

- *Don't think you have to visualize all your data.* Reduce the amount to be included and filter the results to highlight important differences.
- *Don't cram all your data into one visual.* Use two, three, or more charts if it makes understanding easier. This is especially important for complex multi-dimensional data sets where different types of charts are often needed to properly interpret each analysis.
- *Assume viewers are not using a large high-resolution display.* Many viewers will be using mobile devices. It's important to design for and test your visualizations on a wide range of platforms.

ERROR 4: LETTING THE DATA TRIP YOU UP

The main reason for visualizing data is to make it easy to-understand. Unfortunately, errors in data are very common.

Because of errors, visuals can often look fine when applied to small sample of the data, but become totally unpredictable when used with real sources. Data errors can lead to data analysts spending large amounts of time unnecessarily debugging their work to get visuals to reliably reflect the results.

AVOIDING THE MISTAKE

- *Don't assume your data is without errors.* Spotfire makes it easier to identify errors in data sources. It can't fix those errors, but it will stop you from spending unnecessary time trying to get around charting problems when they arise.
- *If possible, work with the original data source.* A common culprit, the Microsoft Excel spreadsheet, is very widely used so that edited copies of the data can proliferate. Consider that you may be working with an incomplete or out-of-date copy and try to use the original data. Spotfire only needs read access, so making this connection should not be a major issue.
- *Use calculated columns to generate cleaner data for building visualizations without disrupting the spreadsheet.* You can often build great visualizations using imperfect data.

ERROR 5: USING THE WRONG TOOLS

Data visualization tools are not equal. Some concentrate on accessing and analyzing a wide variety of data sources at the expense of producing great looking visuals. Others create stunning charts and dashboards, but only with particular data sets. In addition, there are numerous charting tools that were never really designed for large scale data visualization, but are commonly used for reasons of availability, cost, and expediency.

Using the wrong tools can be a costly mistake both in terms of wasted investment and/or wasted time in learning just what you can't do with the product.

AVOIDING THE MISTAKE

- *Don't rely on Excel.* Microsoft Excel good for collating and filtering data, but for creating visualizations it's nowhere near as easy to use as a data analytics and visualization tool like TIBCO Spotfire.
- *Think about your needs now and into the future.* Look for applications able to work with all the data sources you are likely to want to visualize.
- *Think about how you want to use visualizations.* In addition to static charts, you probably want to create interactive visuals with drill-down capabilities as well as dashboards for monitoring key performance indicators (KPIs).
- *Remember visualizations are only part of the story.* Collaborating on your visualizations may be important. Look for tools able to distribute the charts and dashboards you build.
- *Think about devices.* What devices will your viewers want to use to create and consume visualizations? Would you be better off with a cloud-based tool rather than on-premise software?

IT'S UP TO YOU

Now that you know about the most common data visualization mistakes and how to avoid them, one of the simplest solutions is to use TIBCO Spotfire. But don't just take our word for it, check it out for yourself by signing up for a free trial today.