



Introducing Chart Tamer

An Excel Add-In Product from BonaVista Systems

No other software comes close to *Microsoft Excel* in popularity for analyzing and presenting quantitative information. The charts in Excel, which provide Excel's visual analysis and presentation functionality, however, are confusing, inadequate, and sometimes poorly designed. In some respects this problem actually increased with the release of Excel 2007, because it did little for charts other than add superficial dazzle—visual effects that actually undermine a chart's ability to display information effectively.

Chart Tamer attempts to align Excel's charts with the best practices of data visualization, as taught in the book *Show Me the Numbers* by Stephen Few of [Perceptual Edge](#). It *tames* Excel's charts, which seem to exhibit Microsoft's belief that "more is better", by bringing them into line with the data presentation philosophy that "simple is better." We don't need more choices—we need a few good choices that really work.

What Chart Tamer Does

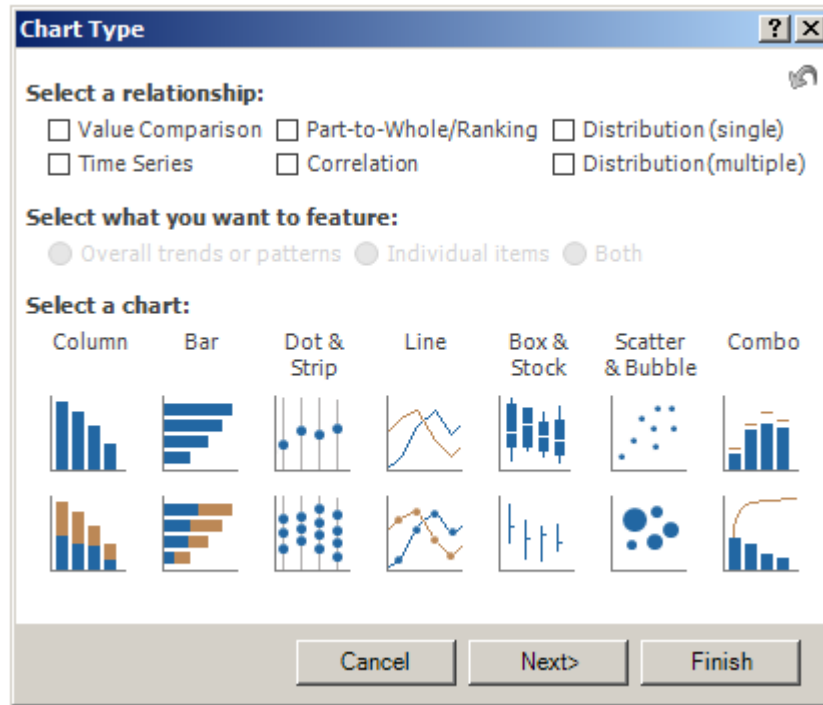
Chart Tamer improves the charts that we produce with Excel in the following ways:

- It limits the list of available charts to the few that are most useful, thereby reducing the complexity of choosing an appropriate chart
- It provides a simple new interface for selecting an appropriate chart, which guides us to the right selection when help is needed without bogging us down when it's not
- It adds three useful charts that aren't currently available in Excel: dot plots, strip plots, and box plots
- It revises the formatting defaults of Excel's charts so that no extra work is necessary to create charts that work effectively
- It provides colors for use in charts and elsewhere that were designed to work especially for data presentation
- It allows us to quickly and easily improve old charts that were created prior to Chart Tamer

Let's briefly tour each of these features.

The Best Charts and When to Use Them

When we use Chart Tamer to create a new chart, the following dialog box makes it easy to choose the most appropriate chart.



Rarely will we ever need a chart other than the 14 that appear in the *Select a chart* section above. If we already know which of these charts will display the data best, we can click it immediately, without having to wade through any other choices. If we want some guidance, however, we can narrow the list of available charts by making selections in the *Select a relationship* and *Select what you want to feature* sections.

Quantitative Relationships

Charts that are used for displaying quantitative information always feature one or more relationships between the items that appear in the chart. Particular types of charts are designed to feature particular relationships in an attempt to tell a particular story. If we know the relationship that we want to feature, we can follow Chart Tamer's suggestions to narrow the list of available charts. Here's a list of relationships that we'll need to feature:

- Value comparison
- Part-to-whole and/or ranking
- Time series
- Correlation
- Distribution (single)
- Distribution (multiple)

Value Comparison

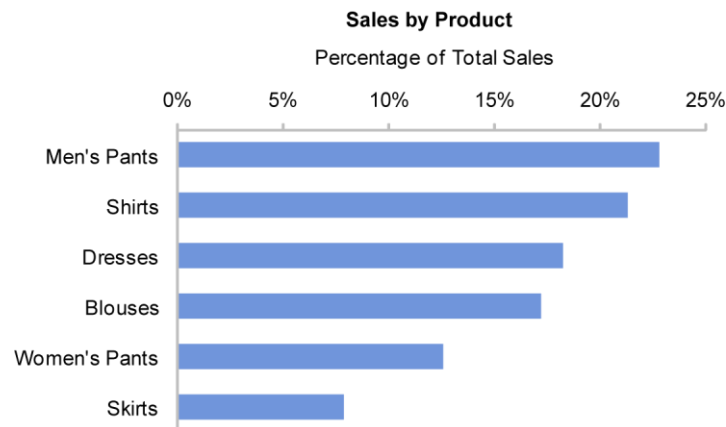
All of these relationships involve comparisons, but the one that Chart Tamer calls a *Value comparison* features only one simple relationship among the values that appear in a chart: differences in their magnitudes. On those occasions when we don't want to feature any other

relationship between the values (for example, how they relate to one another in terms of rank from high to low or low to high), then *value comparison* is the right choice.



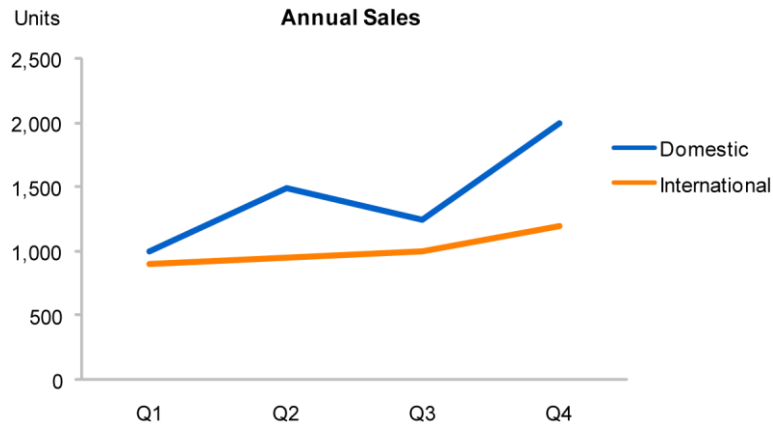
Part-to-Whole / Ranking

At times, the relationship that we want to feature is how a set of values contributes to some whole, such as the regional parts of total sales. This is what's called a *part-to-whole* relationship. Charts that feature this relationship must make it easy for us to compare the parts that make up the whole to one another. People usually display part-to-whole relationships with pie charts, but pie charts are intentionally missing from our list, because they don't actually make it easy to compare the parts of something to one another. It's difficult to compare the slices to one another, because comparing two-dimensional areas isn't something that visual perception evolved to do very well. As an alternative to pie charts, bar charts display part-to-whole relationships much more effectively. When we want to feature and compare parts of a whole, it is almost always useful to arrange the parts in ranked order, either from highest to lowest or vice versa. Even when the items in a chart don't represent some whole (that is, they don't sum to 100%), we often want to feature how those values relate to one another by order of magnitude—what's called a *ranking* relationship. In both cases, *part-to-whole / ranking* is the relationship that we want to feature.



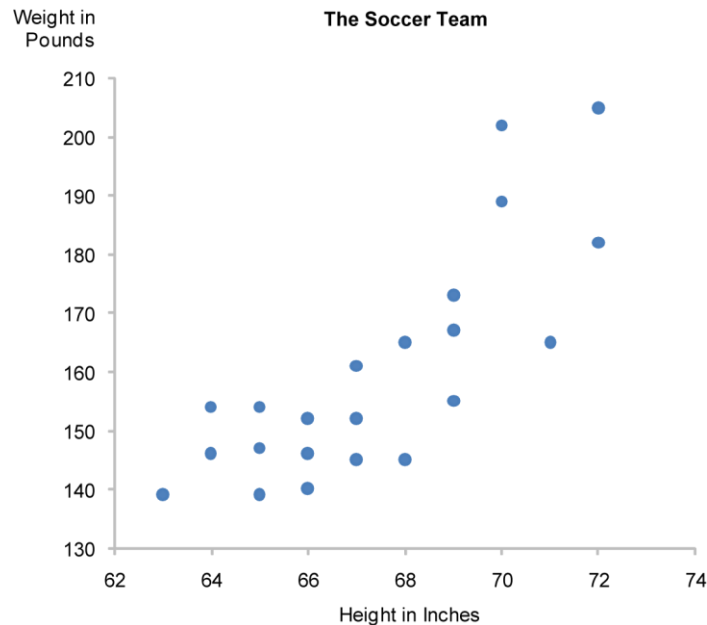
Time Series

No relationship is featured more often in charts than the way values change through time—what's called a *time series* relationship. According to a study that was done a few years ago, 70% of all graphs feature change through time. If this is your objective, *time series* is the appropriate choice.



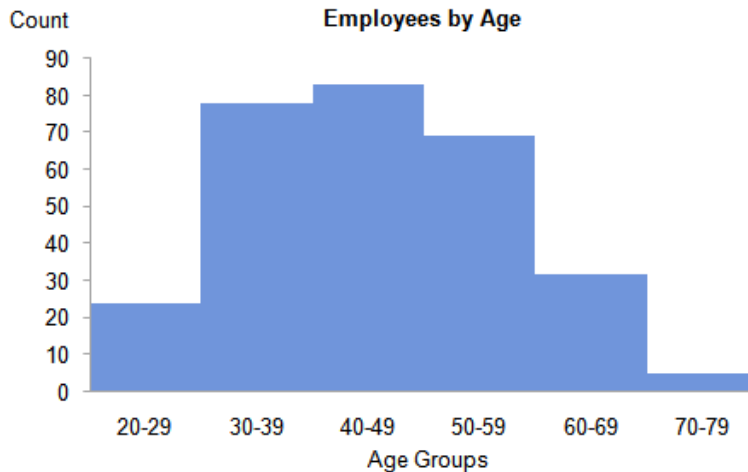
Correlation

Sometimes the relationship that we need to feature is the way two sets of quantitative values, each of which measures a different attribute of something, relate to one another—that is, how they co-relate or correlate. For example, we might want to display a potential correlation between a group of people’s heights and their weights? Scatter charts (also called scatter plots) were originally created to display these relationships. When this is the relationship we need to feature, *correlation* is the appropriate choice.



Distribution (Single)

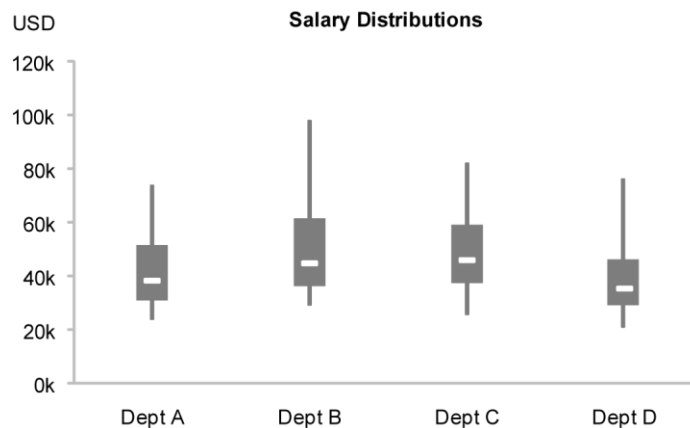
Sometimes the relationship that we need to feature in a set of values is the quantitative range that they are distributed across from the lowest to the highest values, and where the individual items fall within that range. For instance, we might want to display a count of the number of people in our organization by age groups such as 20-29, 30-39, 40-49 years old and so on. This is called a *distribution* relationship. When we feature the distribution of a single set of values (for example, all employees in the company by age), *distribution (single)* is the appropriate choice.



Notice that in this example of a distribution display using a column chart the space that Excel usually places between the columns (that is, what Excel calls the gap width) has been eliminated entirely (the gap width has been reduced to 0%). When column charts are used to display distributions, they're called histograms, and conventionally in histograms the columns touch one another with no space in between. Because we selected "Distribution (single)" from the "Select a relationship" section of the Chart Type dialog box and then chose a Column chart, Chart Tamer automatically eliminated the space between the bars, saving us that extra step.

Distribution (Multiple)

On those occasions when we need to feature and compare multiple distributions (for example, the distributions of employee salaries in four different departments), as opposed to featuring the distribution of a single set of values, *distribution (multiple)* is the right choice. The graph below is called a *box plot*. You might not be familiar with it, but don't worry. It's easy to use and is very effective for multiple distributions.



When we click on one of these relationships to select it, all of the other relationships that cannot be featured in a chart become unavailable, as do all of the choices in the *Select what you want to feature* and *Select a chart* sections that are not appropriate for the relationship we've chosen. In this manner our choices are narrowed, which simplifies the chart selection task.

What We Want to Feature

The list of available charts can be narrowed, often down to a single chart, by making picking items in the "Select a relationship" and "Select what you want to feature" sections. Some charts do a good job of featuring overall trends and patterns, some effectively feature individual items, and some provide a compromise that does a fairly good job of both. For instance, when showing

how something has changed through time (a time series relationship), we often want to feature the overall trend of change or patterns of change, so we would choose *Overall trends and patterns*. Doing this would cause the list of charts to be narrowed to a line chart, because nothing features trends and patterns in change through time better than a line chart. On the other hand, when we display actual expenses compared to the budget by month, and want to make it as easy as possible to compare actual to budgeted expenses in individual months, rather how they differ overall, we would choose *Individual items*, which would narrow the list of available charts to the column chart alone. Given the same situation, however, if we wanted to make it easy to see overall trends and patterns and to focus on individual items as well, we would choose *Both of above*, and the list of available charts would be narrowed to the line chart with data points to mark the individual values.

If you follow the guidance that Chart Tamer provides for awhile, before you know it you won't need any help to select the best chart, because the best choices will become obvious to through experience. In this way Chart Tamer can actually help you learn best practices of data presentation.

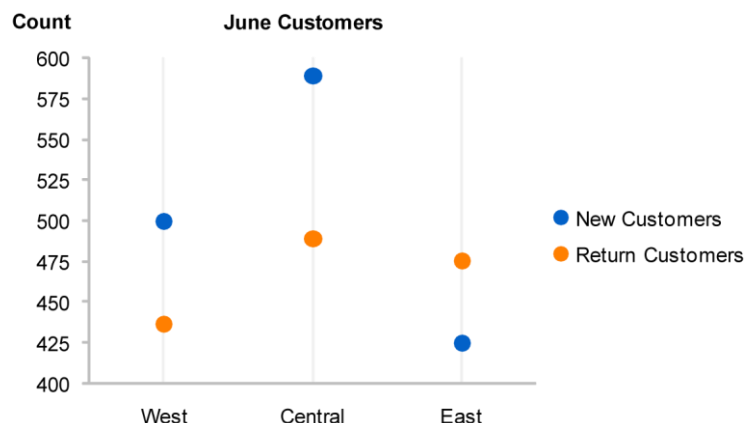
New Charts

Chart Tamer provides three charts that aren't currently found in Excel, because they are incredibly useful:

- Dot plot
- Strip plot
- Box plot

Dot Plots

As mentioned previously, column charts do a wonderful job of featuring individual values in a way that makes them easy to compare. When represented as columns of varying heights, we compare values by comparing their heights. This only works, however, if the chart's quantitative scale includes zero and it serves as the baseline for the columns. There are times, however, when we want to make magnitude comparisons easy, but we don't want the quantitative scale to go all the way down to zero because all the values are relatively close in magnitude and significantly higher or lower than zero. In cases like these, because columns require a baseline of zero, dot plots come in handy as a useful alternative. Here's a simple example of a dot plot:

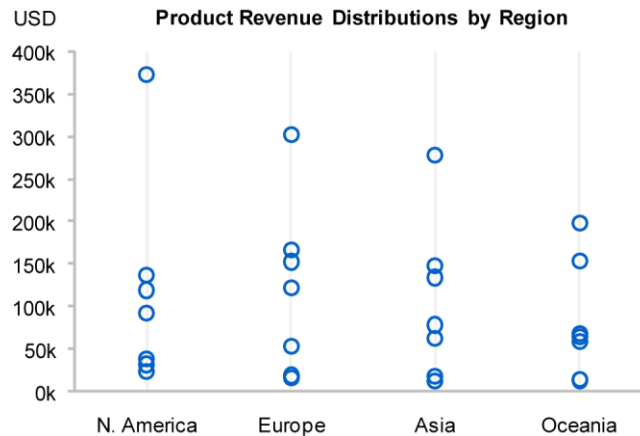


Instead of columns, dot plots place a dot (that is, a marker of some type, such as a small circle or square) at the position where each of the columns would have ended. Because we are no longer comparing the heights columns, it isn't necessary to include zero in the scale, and thus our

problem is solved. Even when we don't need to narrow the scale such that it no longer includes zero, dot plots can be used as an alternative to a column chart.

Strip Plots

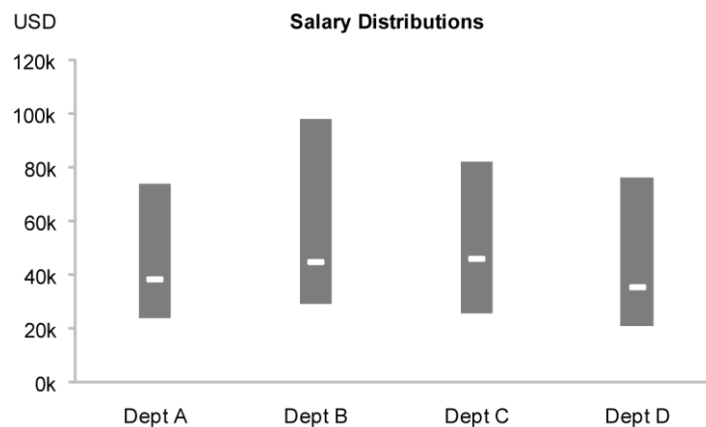
A strip plot looks a bit like a dot plot but it's used for a completely different purpose: to display distributions. The example below shows the distribution of eight product's sales revenues in four sales regions.



As you can imagine, a strip plot would not be an effective way to display the distribution of product revenues if there were 100 products, but when a distribution consists of a few values only that should be displayed in a way that allows us to see the position of each value, a strip plot provides a nice solution. Strip plots can be used to display a single distribution or multiple distributions when each consists of the same set of items, as in the example above. In cases when several values in a distribution are precisely the same, however, which would cause those data points to appear in exactly the same position and some to be hidden, the strip plot solution suffers, but the box plot, which will look at next, does the job nicely.

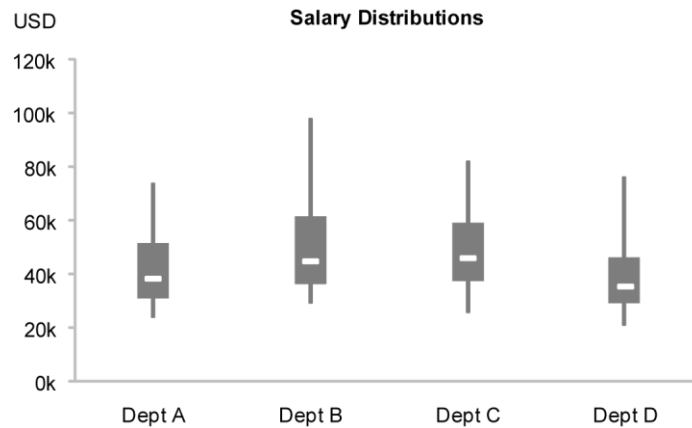
Box Plots

When we want to compare multiple distributions but have too many values to show in a strip plot and also want to clearly feature attributes such as a measure of its center (usually the median or mean), nothing does this better than a box plot. Here's an example of a simple box plot:

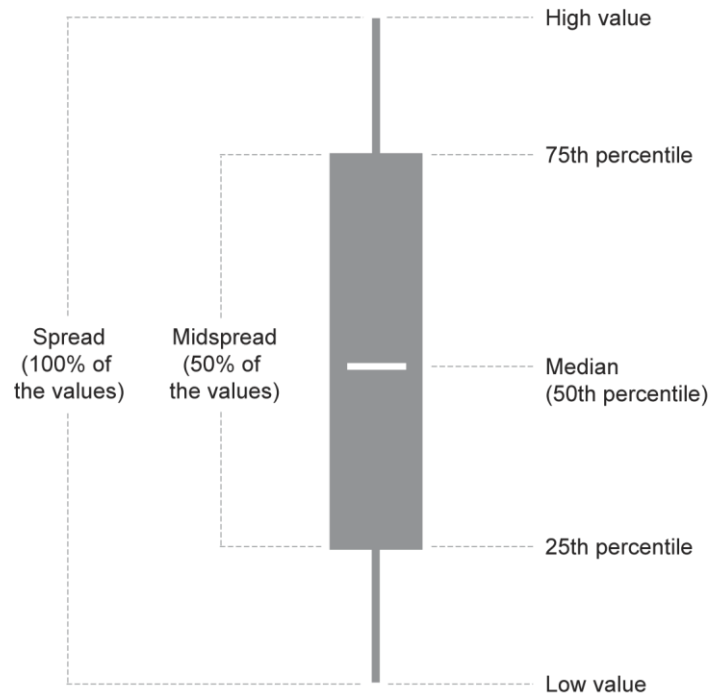


This box plot makes it easy to compare how salaries are distributed among four separate departments. The top and bottom of the gray rectangles in this example represent the highest

and lowest salaries in each department and the center marks represent median salaries. This is called a three-value box plot, but even more information can be conveyed using a five-value box plot, shown here:



The addition of the line provides more information about how salaries are distributed. Box plots of this type can be used to represent various measures of center, usually medians or means, and various measures of dispersion around the center, such as the percentiles and standard deviations. The diagram below explains how a box plot that uses the median as the measure of average, and the 75th and 25th percentiles as the measure of dispersion around the median salary, and highest and lowest salaries as the ends of the range, would be read.



(Note: The values that are displayed in box plots (center, high, low, etc.) may be based on various measures. For instance, even though we've used the ends of the lines (also called "whiskers") to represent the highest and lowest values in the data set in the example above, you may assign any of the other appropriate values that you wish to the ends of the lines, such as 1.5 times the inter-quartile range rather than the highest or lowest values when outliers exist.)

Putting It All Together

The following table shows how each of the quantitative relationships and the different aspects of the data we want to feature are associated with specific charts by Chart Tamer.

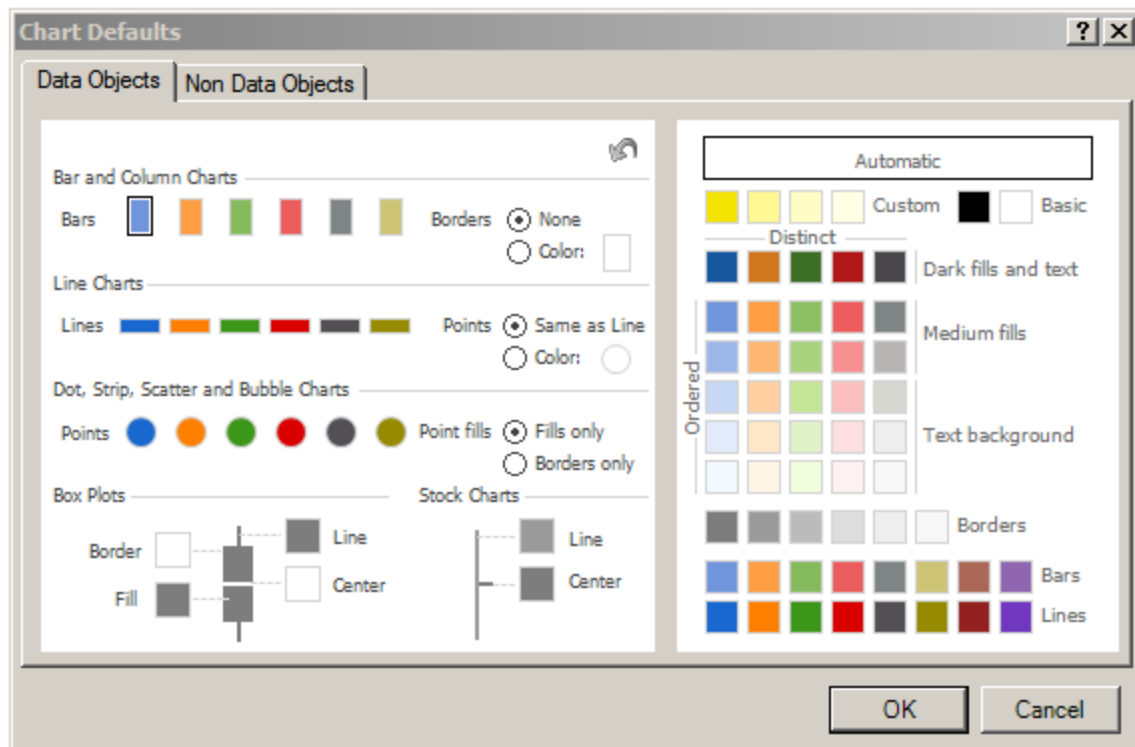
Relationship	What should be featured	Chart
Value comparison	Individual items	<p>Column chart</p> <p>Stacked column chart (when individual columns are broken into parts of a whole)</p> <p>Bar chart</p> <p>Stacked bar chart (when individual bars are broken into parts of a whole)</p> <p>Dot plot (especially when we want to narrow the quantitative scale and exclude zero)</p> <p>Column/Point combo chart (especially when you want to feature one set of values—the values in the columns—more than others)</p>
Part-to-whole / Ranking	Individual items	<p>Column chart</p> <p>Stacked column chart (when individual columns are broken into parts of a whole)</p> <p>Bar chart</p> <p>Stacked bar chart (when individual bars are broken into parts of a whole)</p> <p>Dot plot (especially when we want to narrow the quantitative scale and exclude zero)</p> <p>Column/Line combo chart</p>
Time series	Overall trends and patterns	Line chart
	Individual items	<p>Column chart</p> <p>Stacked column chart (when individual columns are broken into parts of a whole)</p>
	Both of above	<p>Line chart with data points</p> <p>Column/Line combo chart (especially when you wish to feature individual items in one data set and the overall trend and patterns in another)</p>
Distribution (single)	Overall trends and patterns	Line chart
	Individual items	Strip plot
	Both of above	<p>Line chart with data points</p> <p>Column chart</p> <p>Stacked column chart (when individual columns are broken into parts of a whole)</p> <p>Bar chart</p> <p>Stacked bar chart (when individual bars are broken into parts of a whole)</p>
Distribution (multiple)	Overall trends and patterns	<p>Box plot</p> <p>Stock chart (when the values represent stock prices)</p>
	Individual items	Strip plot (when there are few values and you want to display each)
Correlation	Individual items	<p>Scatter (two variables)</p> <p>Bubble (three variables)</p>

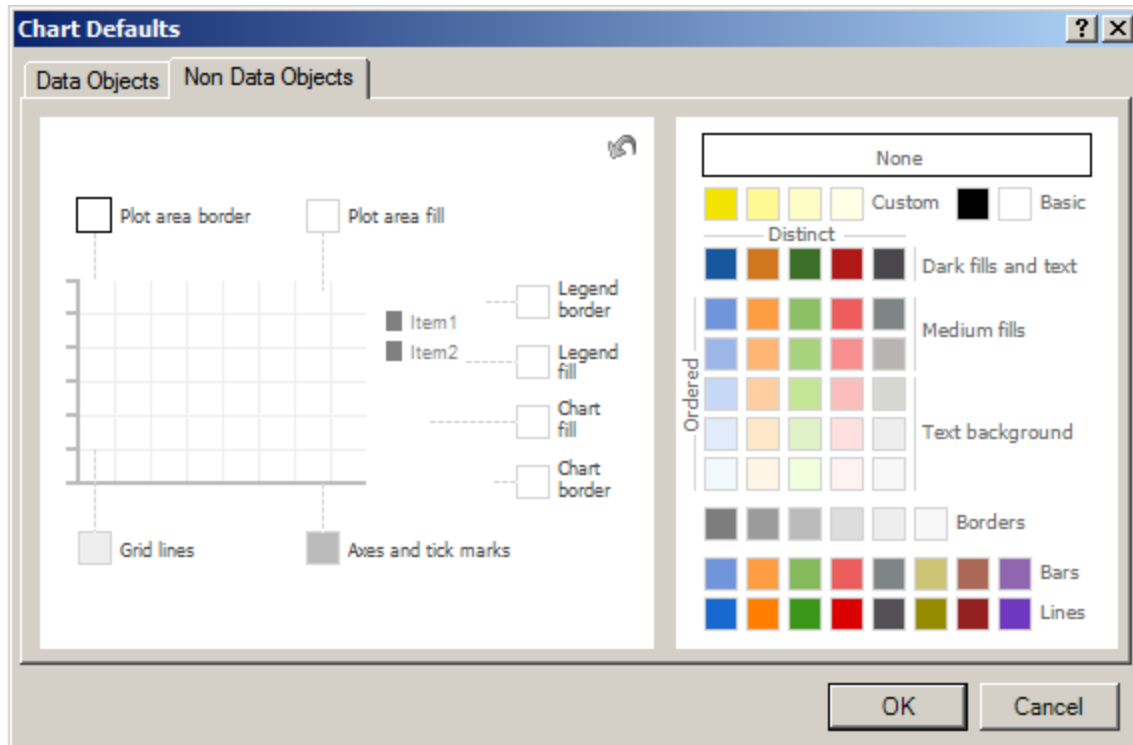
Effective Chart Formatting Defaults

A fundamental principle of effective chart design is that the information should stand out and that everything in the chart that isn't information, such as axis lines, borders around the plot area, and so on, should be just visible enough to do their job, but no more. The message in a chart is contained in the data, after all, not in the non-data components that are there simply to support the display of data. Anything that appears in a chart that isn't data and isn't needed to support the display of data doesn't belong there.

The formatting defaults of Excel's charts fail to give us a good start, making it as easy and efficient as possible to display data effectively without having to adjust a chart's format. We've taken great care in Chart Tamer to set formatting defaults for each of the charts to work as effectively as possible for normal circumstances. For example, axes are light and thin, plot areas don't have borders or fill color so the information stands out clearly against a clean background, legends don't have borders or fill color so they don't draw too much attention to themselves, and grid lines are turned off but when they're turned on they are very light thin lines, to mention a few.

Although the formatting defaults that come with Chart Tamer are well designed, in some cases there are alternatives that would work as well. For instance, light gray grid lines (the default) work well, but for some reason we might want them to be light blue instead. For this reason, we have the ability to change the Chart Tamer's defaults using the Chart Defaults dialog box, which is divided into two tabbed folders: one for setting the defaults of data objects (bars, lines, etc.) and one for setting the defaults of non-data objects (axes, grid lines, etc.).





Effective Chart Colors and How to Use Them

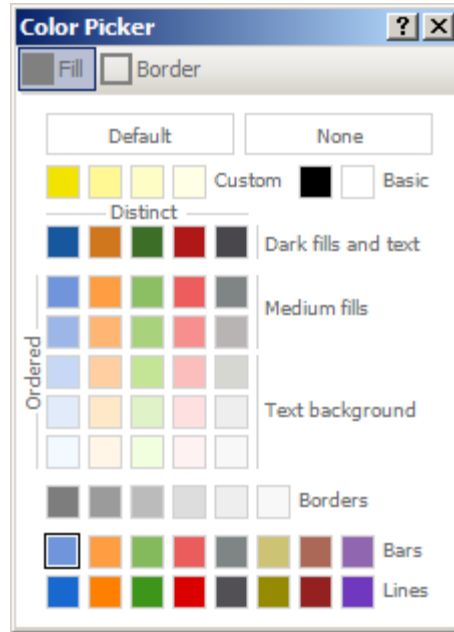
What we perceive as color is nothing more than a specific range of electromagnetic waves—light waves—that enter our eyes either directly from the source that produces them (for example, the sun or a light bulb) or indirectly after reflecting off the surface of an object. We experience some color combinations as beautiful and pleasing and others as ugly and stress-inducing. Achieving the desired effects with color in charts requires a bit of design knowledge and skill, which we've built right into Chart Tamer.

When we present information graphically, such as a year of monthly sales revenues by product in a line chart, we can use colors to distinguish groups of related data or to highlight particular items. For instance, in a line graph that displays actual expenses and the expense budget, we could use a salient color to highlight the actual expenses line, while using a more muted color for the budgeted expenses line. Few aspects of data presentation are as misunderstood and misused as color. Poor use of color can produce an unpleasant display that no one will bother to examine. Good use of color can invite people to explore the information and focus their attention on the most important information in the chart. It's a sad fact, however, that most software products that we use to display information don't guide us toward good color choices. Unfortunately, this is true of Excel, although the colors that were introduced in Excel 2007 are a dramatic improvement over those in earlier versions.

The colors that Chart Tamer provides were carefully selected by Maureen Stone of [StoneSoup Consulting](#), an expert in the use of color for data presentation. Chart Tamer makes it easy to avoid bad uses of color in the following ways:

- It provides color palettes that were carefully selected for their ability to present data effectively, which means that you would have to go out of your way to select poor colors
- It sets default colors for all the components of charts (line colors, bar colors, etc.)
- It simplifies and streamlines the color selection process through its Color Picker
- It teaches us effective color design practices through example

Because versions of Excel prior to Excel 2007 limited the number of colors that can be used to 56, Chart Tamer honors this limit as well for the sake of compatibility. Without this limitation, Chart Tamer could have provided many more palettes of colors from which to choose that would all work equally well, but the truth is, most of us will never need more colors than those that are provided. Here's the Color Picker that we use whenever colors need to be assigned to something in Excel:



The two rows of colors at the bottom labeled “Bars” and “Lines” provide colors that work extremely well for displaying data in charts. Notice that the bar colors are a little lighter than the line colors. This is because large data objects that use fill colors, such as columns and bars, contain a great deal of color and therefore don’t need to be as dark to be easily seen and distinguished as small objects such as lines and data points, which usually need to be a little darker. The first five colors in each of these two rows have been assigned as defaults to bars (the fills row), to lines (the lines row), and to data points (also the lines row). Nevertheless, the final two colors in each of these two rows work as well as the others and are available to use as defaults if we prefer.

Because on occasion we’ll want colors that are a little darker or a little lighter than those in the bottom two rows, several versions of these colors (the first six only) are available in the middle of the dialog box, ranging from dark rows at the top to light rows at the bottom. These have been labeled to suggest some of their primary uses: dark fills and text, medium fills, light fills, and as background colors for text (for example, in the cells of a spreadsheet that contain text). Notice that in any given row, each color is equal in visual weight (saturation and darkness) to its companions. This is important, because usually when we use colors to distinguish different groups of bars in a bar chart, different lines in a line chart, or groups of data points in a scatter chart, we don’t want any one group to stand out more prominently than the others. Each group should look different, but equal in importance. On those occasions when we do want some items to stand out more than others, such as a particular line in a line chart, we could choose a color from one of the darker rows to do this, and use lighter colors for the other lines.

The five rows of colors labeled “Ordered” have been designed to serve another purpose that is useful in charts and elsewhere as well. At times, rather than making different items look different but equal in importance, we want them to look different in a way that suggests a range of value or importance from low to high. For example, if we’re displaying five lines in a line chart, each for a different product, and we want the lines to appear as if they are ranked in order of importance from least to greatest, we could select from one of the columns of colors—blue for instance—and

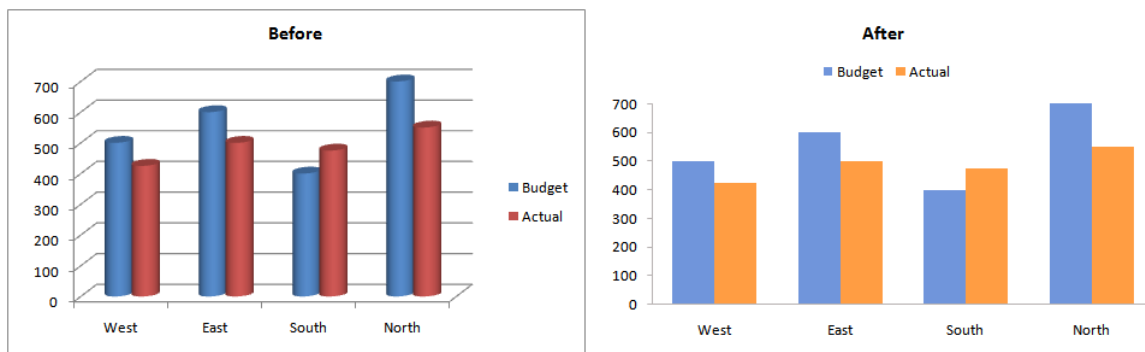
assigned a different shade of blue to each, using the lightest for the least important and the darkest for the most important. Each column of colors in the section labeled *Ordered* was carefully designed to vary in both lightness and saturation in a manner that we naturally perceive as an evenly stepped ramp of colors from low (light, lowly saturated) to high (dark, highly saturated). In addition to using ordered colors in charts, they also come in handy as background fills in cells of a spreadsheet when we want some cells to stand out more (that is, as greater in importance or value) and some to stand out less.

Other colors have been provided for additional purposes, such as shades of gray, which are especially useful for borders (cell borders, chart borders, etc.), black (most useful for text), and a set of four colors that are customizable. The customizable colors are especially handy if your organization has official brand colors that you want to include in charts and elsewhere.

If you're interested in learning more about color and how to use it effectively in charts, you can read a separate document titled [Practical Rules for Using Color in Charts](#) to learn more about this. (The document is located in the folder Start > BonaVista Chart Tamer > Tutorials).

Cleaning Up Old Charts

If you're like most people who use Chart Tamer, you've been using Excel for years and have a huge library of Excel files that were created before Chart Tamer came along. You'd like all your charts to exhibit the effective design practices that Chart Tamer follows, but don't want to recreate those old charts that you still rely on. Fortunately, you won't need to. You can redesign your existing charts simply and quickly using the *Chart Cleaner* feature. If you clean a chart that isn't one of those that Chart Tamer supports (most likely by intention, because it isn't very useful), Chart Cleaner will change it into a chart that will work better and will conform all of its components to Chart Tamer's formatting defaults. It's that easy.



(Note: We owe the idea for this feature to our talented friends at Juice Analytics (www.juiceanalytics.com), who paved the way with an earlier version of this that they gave away for free.)

We hope you find that Chart Tamer helps you get your message across. If your message is worth passing on to others, it's worth doing it well.