

Chart Recipe eBook



by Mynda Treacy

Knowing the best chart for your message is essential if you are to produce effective dashboard reports that clearly and succinctly convey your message.

Contents

Introduction	2
Data Types and Scales.....	2
What’s Your Message	4
Which Chart	5
Nominal Comparisons.....	5
The message:	5
Chart Types	5
Time Series.....	6
The message:	6
Chart Types	6
Ranking.....	7
The message:	7
Chart Types	7
Parts to a Whole	8
The message:	8
Chart Types	8
Deviation or Variance	10
The message:	10
Chart Types	10
Distribution	11
The message:	11
Chart Types	11
Correlation	13
The message:	13
Chart Types	13
A Word on Zero Axis	14
Closing.....	15
References	16
About the Author	16

Introduction

Charts are a useful tool in communicating messages through a picture. That is the shape of the data contains a pattern, trend or exception.

However if you need the reader to be able to get specific values from the data then you might consider putting the data in a table instead, or include a small data table under the chart.

Note: In this guide I will use the term 'chart', but chart and graph are essentially the same thing. Chart is simply the name given by Microsoft for the graphing tools in Excel.

Data Types and Scales

Most 2D charts contain two types of data:

- ❖ Quantitative – the numbers
- ❖ Categorical – the categories those numbers fall into

Without the categories the numbers have no meaning.

The exception to this is a scatter plot which has two quantitative scales.

Different types of categories are:

- ❖ Regions - north, south, east and west or Asia, Americas, Europe etc.
- ❖ Departments - sales, marketing, operations etc.,
- ❖ Time - days, weeks, months, quarters, years etc.
- ❖ Bands/Bins – 0-10, 11-20, 21-30 etc.

Charts can contain multiple categories e.g. regions and time.

Categories fall into one of 3 types:

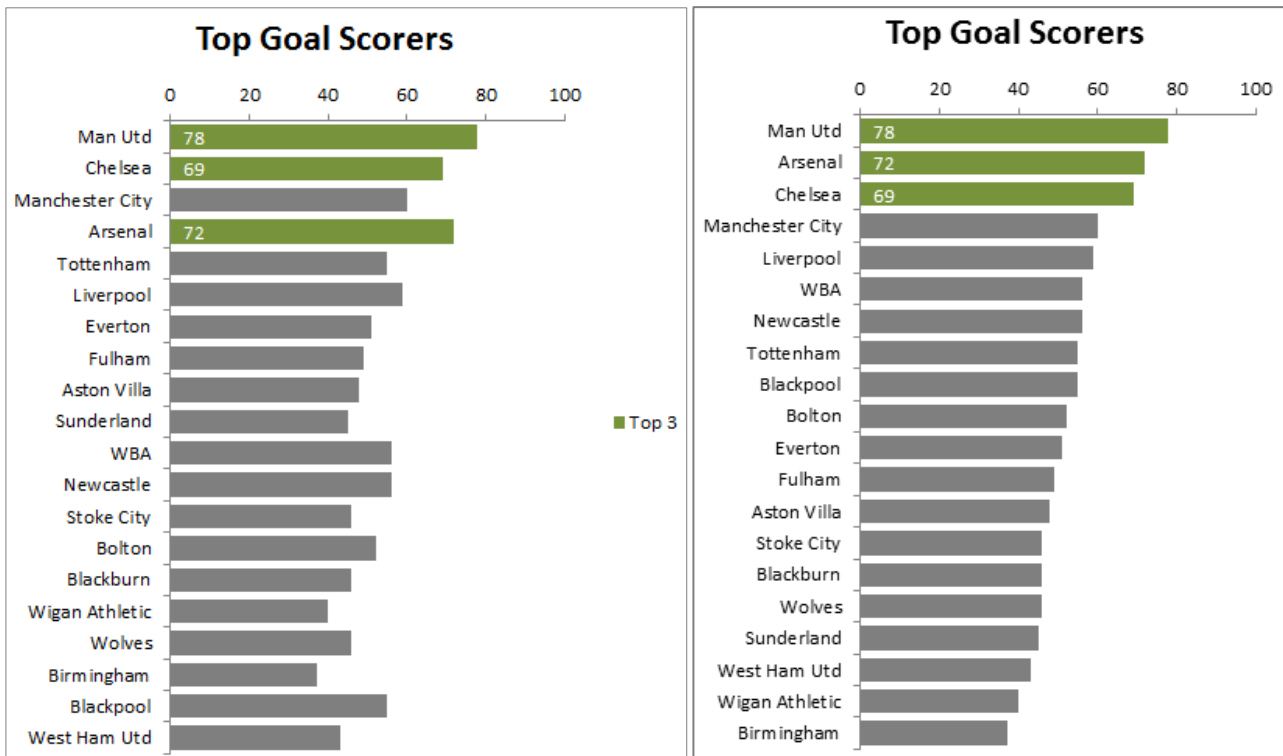
- ❖ Nominal – think 'name'.
- ❖ Ordinal – think 'order'.
- ❖ Interval or Ratio – think 'groups'.

Nominal scales have no specific order, however you might order them based on convention e.g. North, South, East and West, or you might always list your departments in alphabetical order, but the order itself does not communicate any message in the data.

If you were to sort them based on the numbers in the chart e.g. highest to lowest you would give them an order and convey a message about that data in doing so.

Let's take the two bar charts below as an example. The first chart is sorted based on the team's position in the league table, whereas the second chart is sorted based on the values in the bars. By sorting the results we are giving the nominal scale order.

Notice also that the second chart doesn't need the legend as it is obvious that the highlighted bars represent the top 3 goal scorers.



Ordinal scales are the opposite, in that they do have an intrinsic order. You must therefore present them in their order because any other way would be confusing. For example, you would never say medium, large, small or good, very bad, very good, average.... The reader would think you came from another planet!

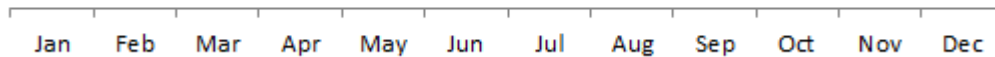
Interval or Ratio scales are the result of grouping your data into equal intervals, for example delivery time 1-3 days, 4-6 days, 7-9 days and so on.

A common mistake when calculating the intervals is to have unequal intervals. You must ensure that the intervals you choose don't result in misinformation.

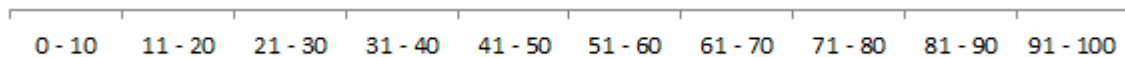
Nominal Scale



Ordinal Scale



Ratio Scale



What's Your Message

According to Stephen Few there are seven types of quantitative messages. That is, what is it that you want to say with your chart?

- ❖ Nominal comparison – comparing departments, regions etc.
- ❖ Time-series – trend over time
- ❖ Ranking – largest to smallest, smallest to largest
- ❖ Parts to a whole – actual vs budget/target, regional share of overall revenues
- ❖ Deviation – data compared to the norm or expected result expressed as the difference
- ❖ Frequency distribution – counts of data per interval or range
- ❖ Correlation – comparison of two paired sets of measures to understand relationship between measures

A chart can contain one or more of these messages.

Once you know what your message is you can figure out which chart is going to be the best fit. Let's take a look at some suitable charts for the different types of quantitative messages above.

Which Chart

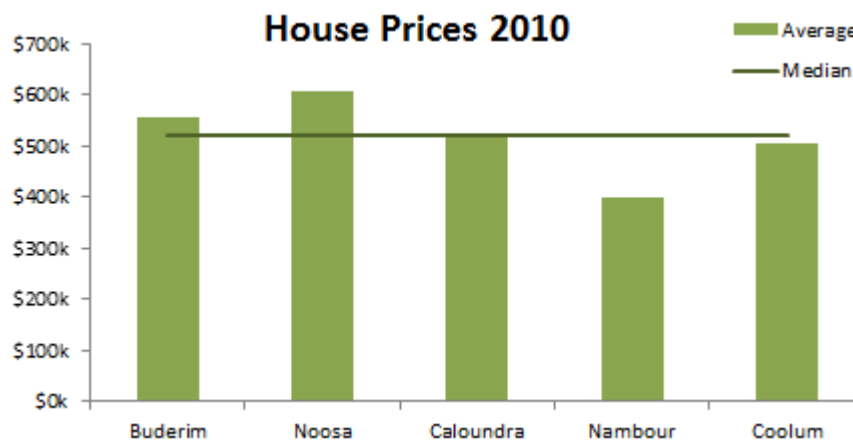
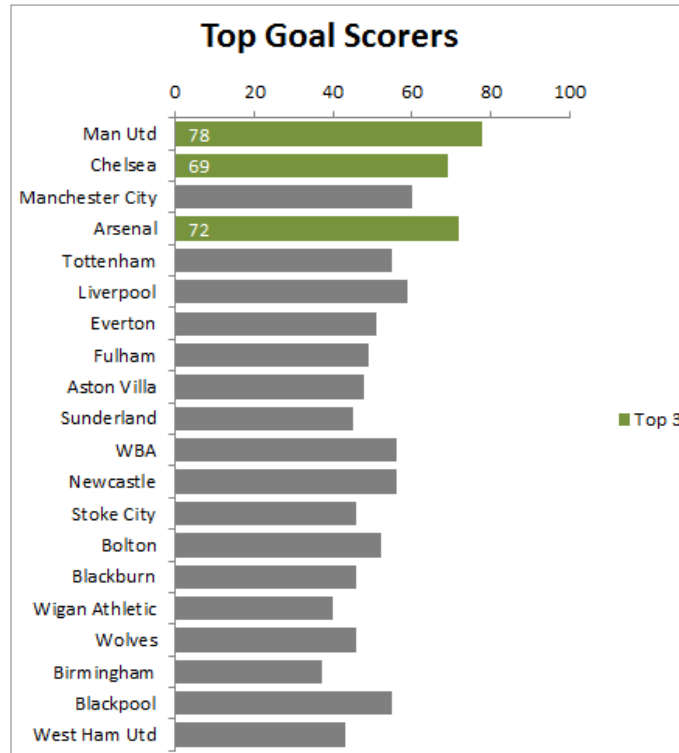
Nominal Comparisons

The message:

Comparing values across categories

Chart Types

Column or Bar charts:



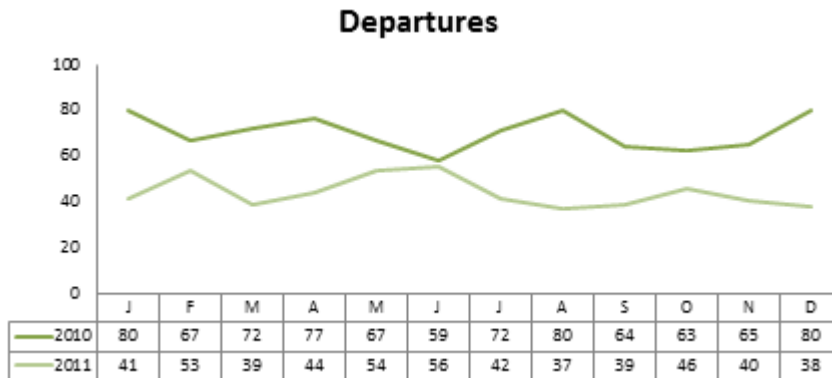
Time Series

The message:

Displaying the pattern of one or more measures over time.

Chart Types

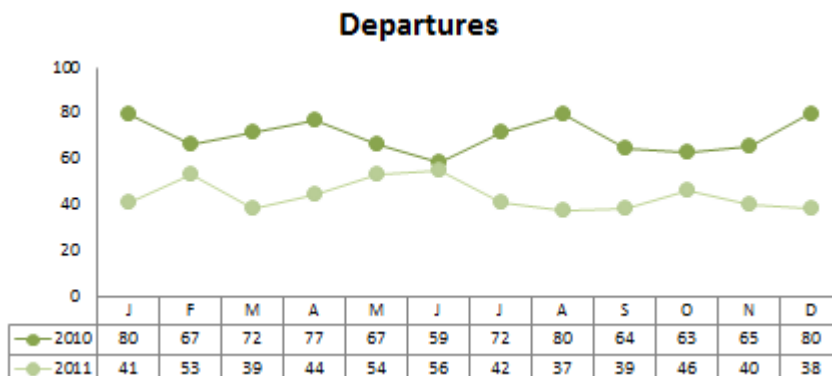
Lines focus on the overall pattern:



Bars focus on the individual values:



Best of both worlds: Dots to emphasise individual values (similar to a column chart) and fine lines to convey overall pattern like a line chart.



Step Charts: these are useful for displaying how the levels in your data increase, remain constant or decrease over time:



Since lines connect the individual values they emphasise the shape of the data as it moves from one value to the next. You should therefore only use lines along an interval scale. With nominal and ordinal scales the values are not closely related and so linking them with a line is implying a relationship that doesn't exist. Instead you should use bars or dots/markers.

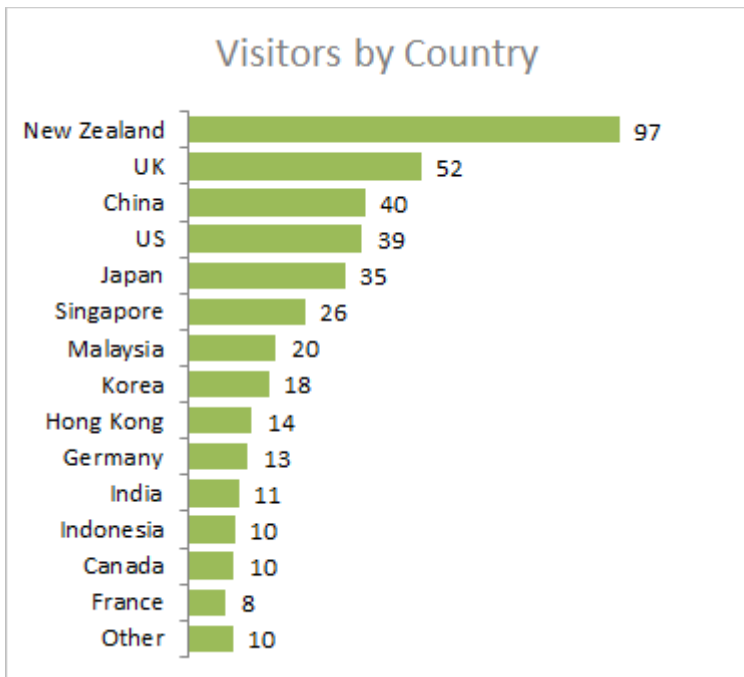
Ranking

The message:

Highlight high values sort in descending and to highlight low values sort in ascending order.

Chart Types

Bar charts, in-cell charts or Sparklines:



Parts to a Whole

The message:

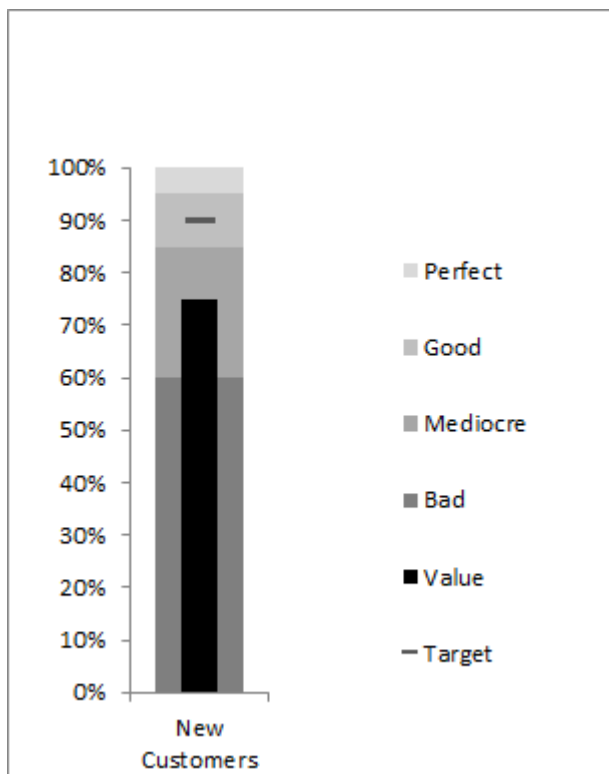
Each value's portion compared to the whole.

Chart Types

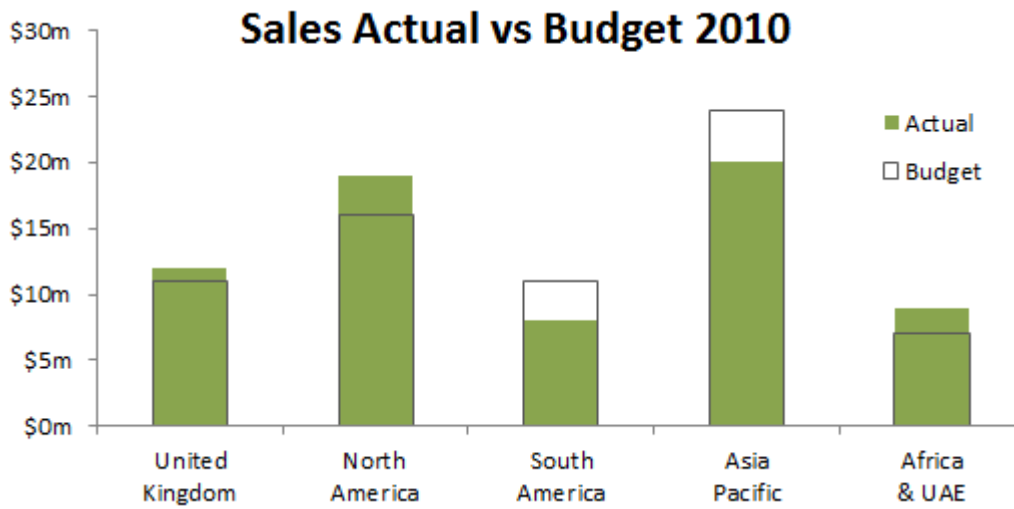
Bar charts or column charts:



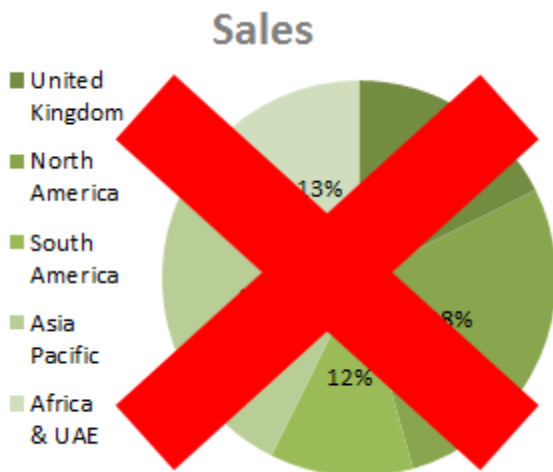
Bullet charts:



Occasionally stacked bars:



But only pies if there are 3 or less items/segments:



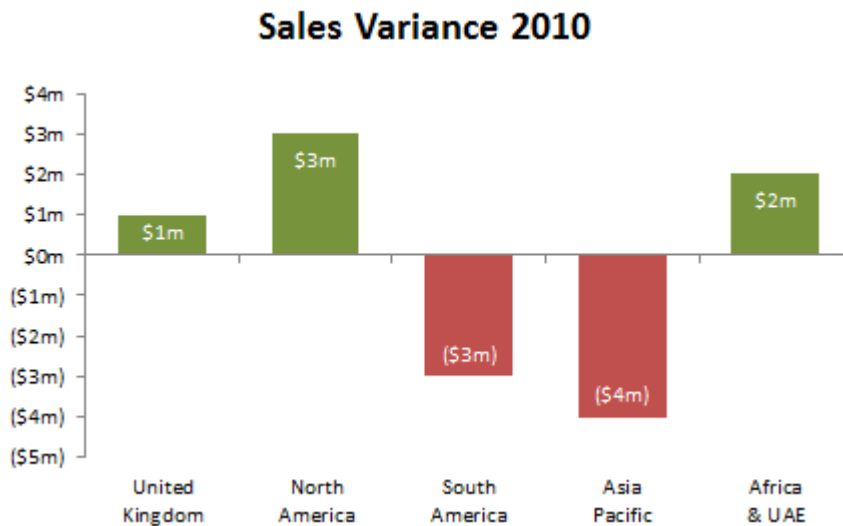
Deviation or Variance

The message:

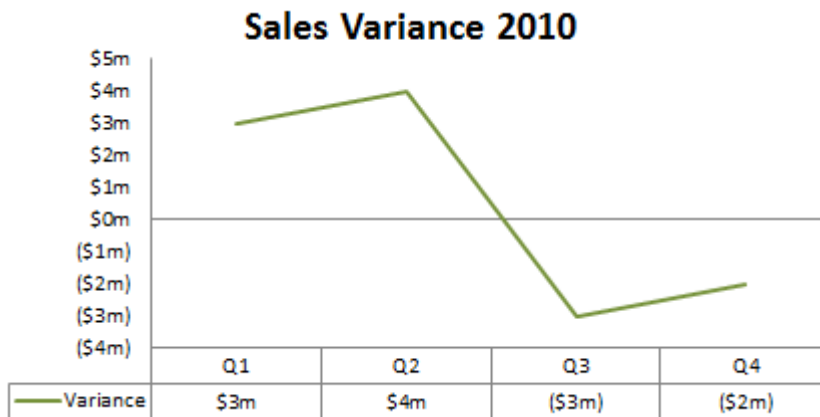
How your values compare to an expected value or target. For example; budget vs actual, compared to a tolerance etc.

Chart Types

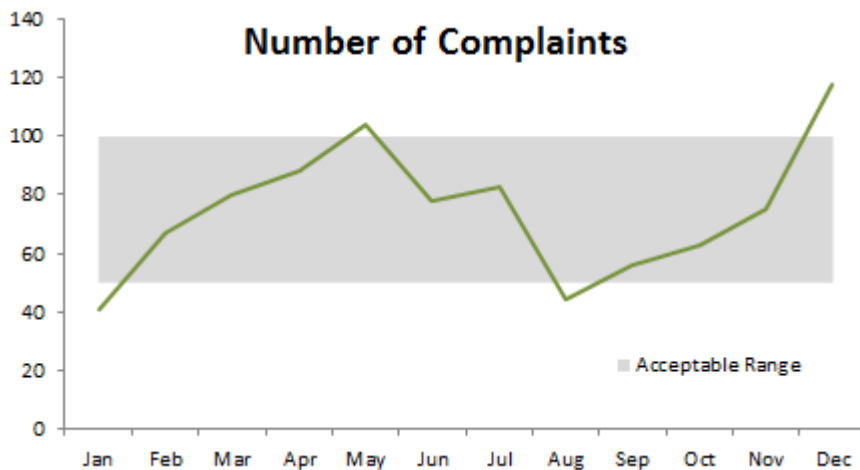
Column or bars to emphasise individual variances:



Line charts allow you to see the variance over time:



Where your target is a range you can use a line or column chart compared to a range or tolerance:



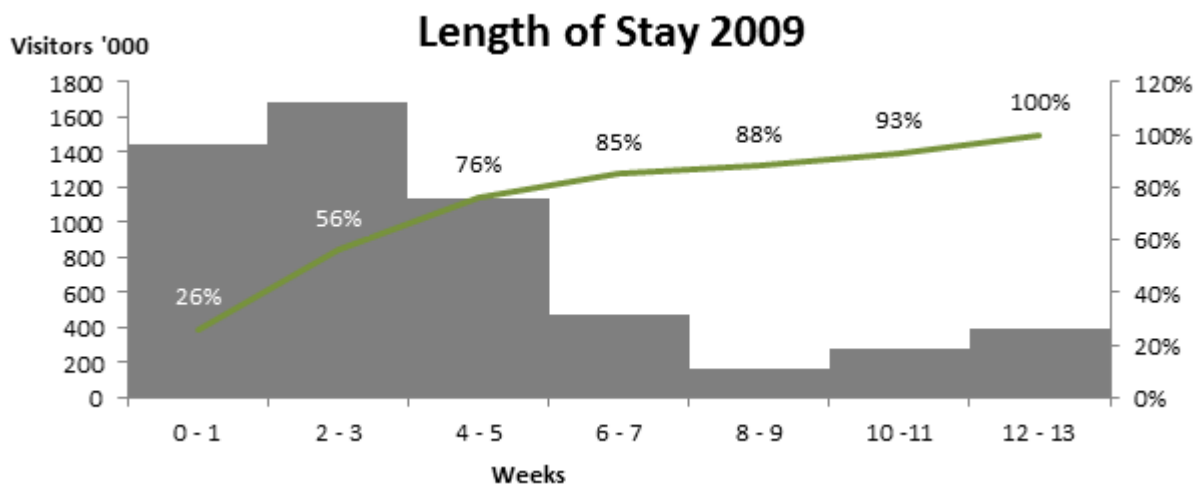
Distribution

The message:

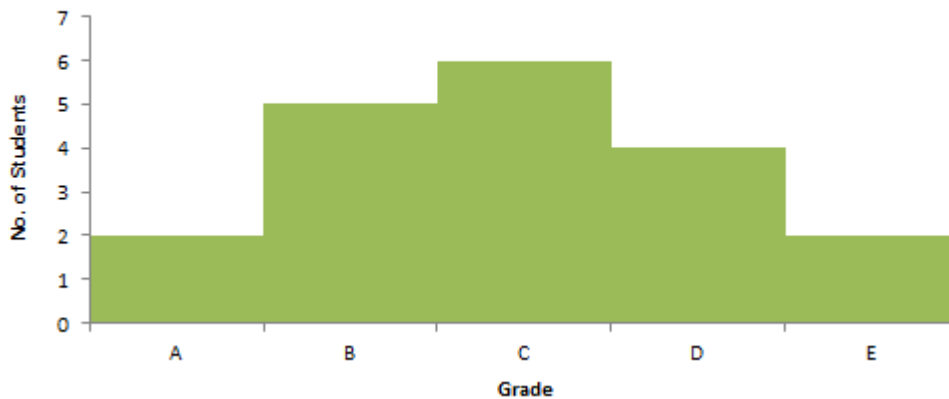
Counts grouped into intervals.

Chart Types

Histogram – note: since a Histogram is a combination of columns and bins you simply use a column chart with 0% gap to plot your data that you’ve grouped into bins. See note on bin sizes below.



Grade Distribution



Take care when choosing your bin size that you don't cover up any anomalies in the data by making your bins too big. Experiment with bin sizes, small bins and larger bins to check that the message is consistent in both.

Take this example from Lealand Wilkinson's 'The Grammar of Graphics' paper:

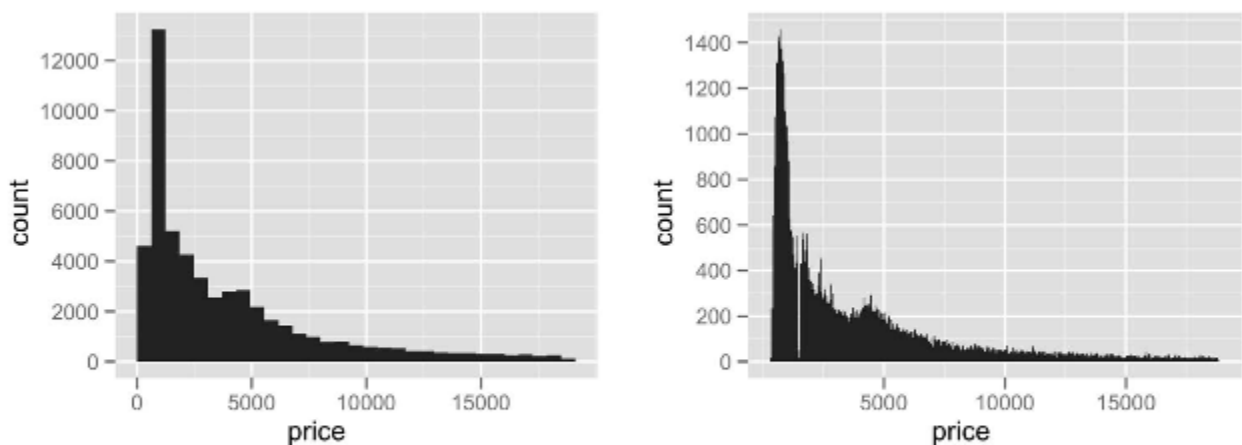


Figure 13. Two histograms of diamond price produced by the histogram geom. (Left) Default bin width, 30 bins. (Right) Custom \$50 bin width reveals missing data.

You can also use line charts with interval scales to emphasise the overall shape of the data.

Box Plots (box and whisker chart) are another alternative to compare distributions but these are not as easily interpreted to the untrained user.

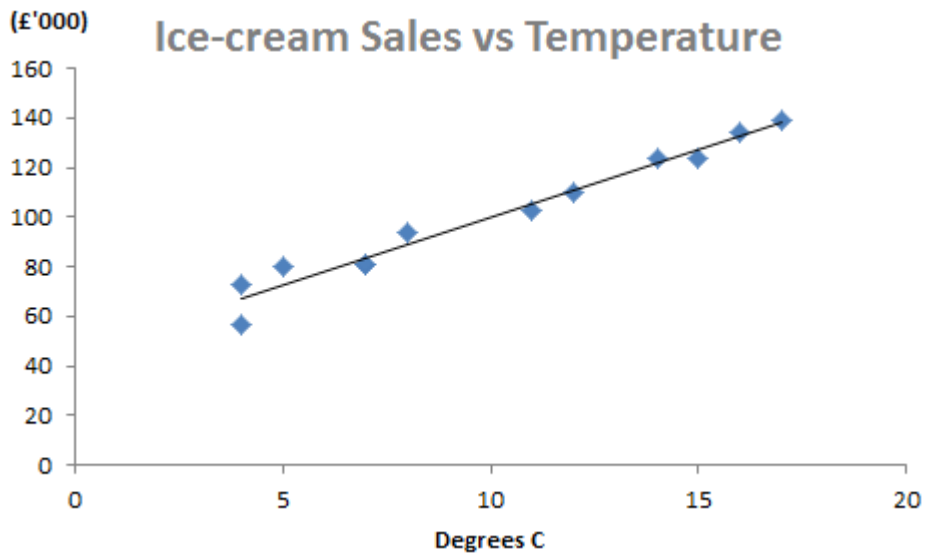
Correlation

The message:

Detecting a relationship between two separate values; for example if one goes down does the other go up and vice versa. Or in the case of ice-cream sales and temperature, when one goes up so does the other.

Chart Types

Scatter chart with a trend line:

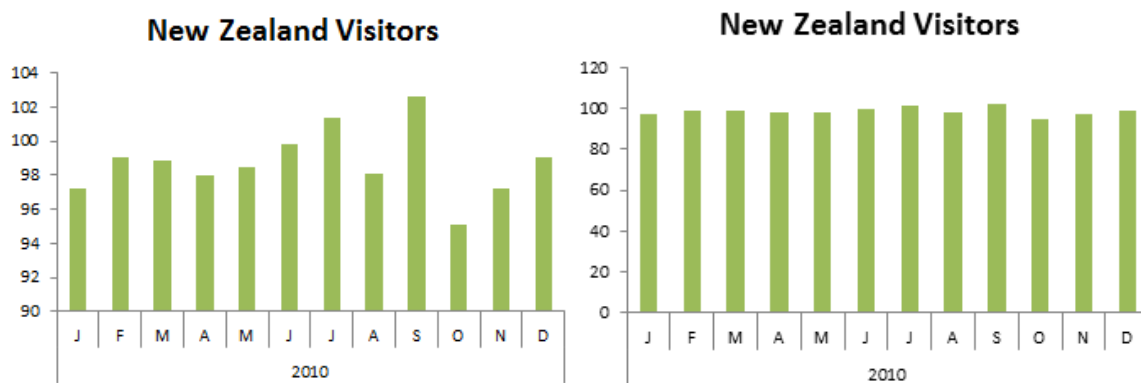


A Word on Zero Axis

Remember the rule that the axis on a column or bar chart must always start at zero so that you don't mislead the reader.

The reason it is misleading with bar and column charts is because we instinctively make judgements about the length of each bar compared to the others. If we start our bar above zero we overemphasise the difference from one bar to the next.

Whilst this looks dramatic in a chart it essentially misleads the reader. Take these two charts below that plot the same data. In the first chart the visitors in September appear to be double that of October, but if we look at the second chart where the axis starts at zero we can see that the difference is actually much smaller.



A commonly accepted alternative when starting the axis above zero is to use markers instead of columns:



Points highlight the individual values rather than the height or length of a bar, and the reader references the quantitative scale to make judgements about the differences between the values.

I still prefer to start the axis at zero and use a column or bar chart as I think they're quicker and easier to read, but if you have a valid reason for not starting your axis at zero, then using points or markers is an accepted alternative.

Closing

So now that you have a recipe book of charts you can go ahead and add your data ingredients and see what delicacies you can come up with, but don't forget that when it comes to formatting, less is more.

When choosing your chart consider the type of data you have but also the message you want to convey. If you have multiple messages consider building multiple charts or making the charts dynamic so the user can choose the view.

Of course you aren't limited to the charts in this guide but following these rules is a good starting point and bear in mind that sometimes your dashboards might need a sprinkling of data tables too.

References

Stephen Few – Communicating Numbers

Lealand Wilkinson – The Grammar of Graphics

About the Author



Mynda Treacy is co-founder of My Online Training Hub and head Excel Geek.

She has worked with Excel for over 20 years, cutting her teeth as a management accountant in global investment banks in London.

Today she teaches people how to harness Excel to produce amazing, interactive dashboard reports.

She has been awarded the Microsoft Excel MVP award every year since 2014 for her contribution to the Excel community.

You can learn from her at www.MyOnlineTrainingHub.com