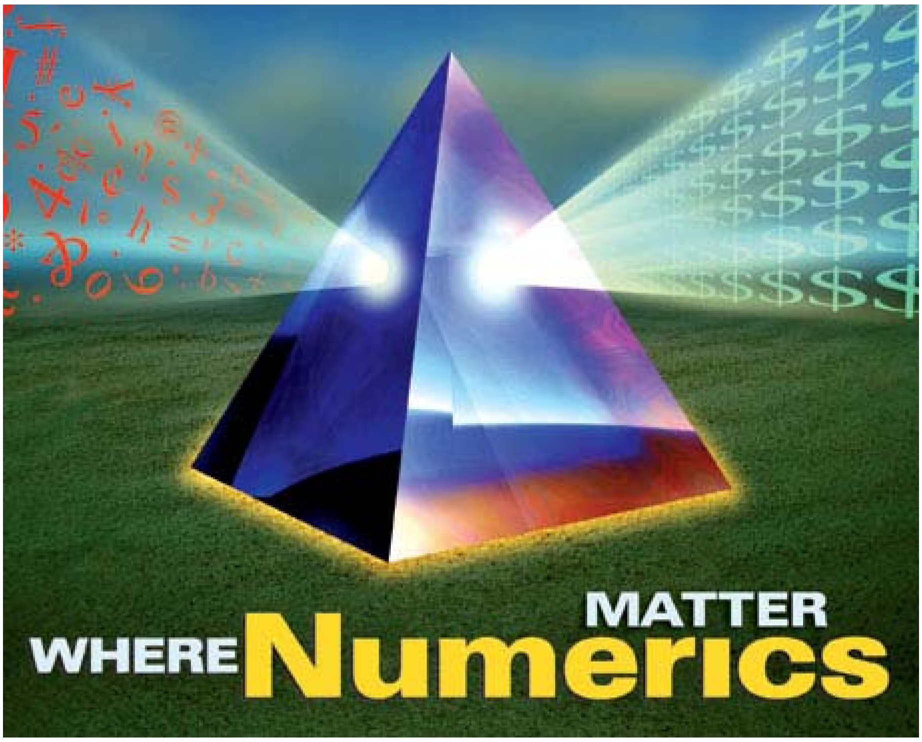

Get The Picture: Visualizing Financial Data – part 1

by Jeremy Walton

Turning numbers into pictures is usually the easiest way of finding out what they mean. We're all familiar with the display of – for example – stock prices as they change with time, and have become accustomed to quickly making decisions based on what the picture is telling us. The reason for this isn't hard to find: images are easier to understand than numbers because our eyes and brains are good at working together to discern visual relationships.

In addition, a well-chosen image can represent a vast amount of numerical data – turning it into a picture is one way of reducing the amount of space that it occupies on the page. Moreover, if it's possible to interact with the picture – by, say, changing the data, or the resolution of the display – it becomes possible to "drill down" into the dataset by focussing on a region of interest and refining the level of detail. If the display is an interactive three-dimensional one, the viewer can fly through the data by changing their viewpoint, searching for interesting features. Finally, if the user has access to the data source, they can perform "what-if" experiments, making predictions of future behavior based on past trends and relationships between different parts of the data.

The art of creating figures or images from numerical data is called *visualization*, and software packages for performing data visualization have been with us for some time. Other applications have incorporated some visualization features alongside other functionality. For example, Microsoft® Excel includes the facility for producing charts and plots from the contents of a spreadsheet via its so-called chart wizard. Features of the plot such as axis labels and data ranges can be edited interactively; in addition, because the plot is linked to the data in the spreadsheet, changes in the data are immediately reflected in the plot, which facilitates the "what-if" analysis mentioned above.

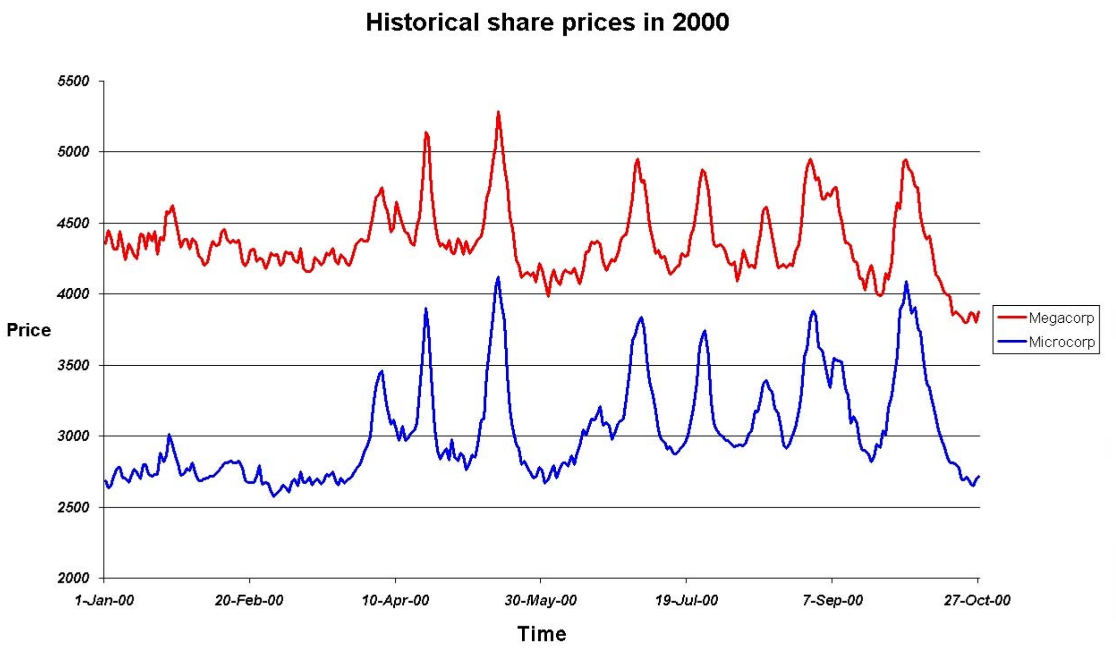

Figure 1 *Using Microsoft® Excel to display the share prices for two companies throughout part of 2000. All aspects of the plot – colors, line widths, label text, fonts and placement – can be edited interactively. The plot is linked to the data in the spreadsheet; any changes in the latter are immediately reflected in the former.*

Plots and charts are good for displaying one variable as a function of another – for example, share price vs. time. A number of share prices – or other variables such as the trading volume for each company – can be displayed on the same plot, allowing comparisons and predictions to be made. Figure 1 shows an Excel plot of share prices for two companies over much of 2000; the trends and similarities in the two datasets can easily be discerned.

However, some financial data are more complicated, and requires other techniques. For example, calculations in the theory of options on swap agreements (so-called "swaptions") using a model like Black-Scholes produce values for the option *volatility* as a function of the underlying *maturity* of the swap and the *expiry* time on the option. Here, the volatility depends on two variables (maturity and expiry) and this can suggest a helpful distinction that can be drawn here between these two *independent* variables and the *dependent* variable (the volatility). Data like this can be displayed using contour or surface plots with the independent variables being measured along axes in two dimensions.

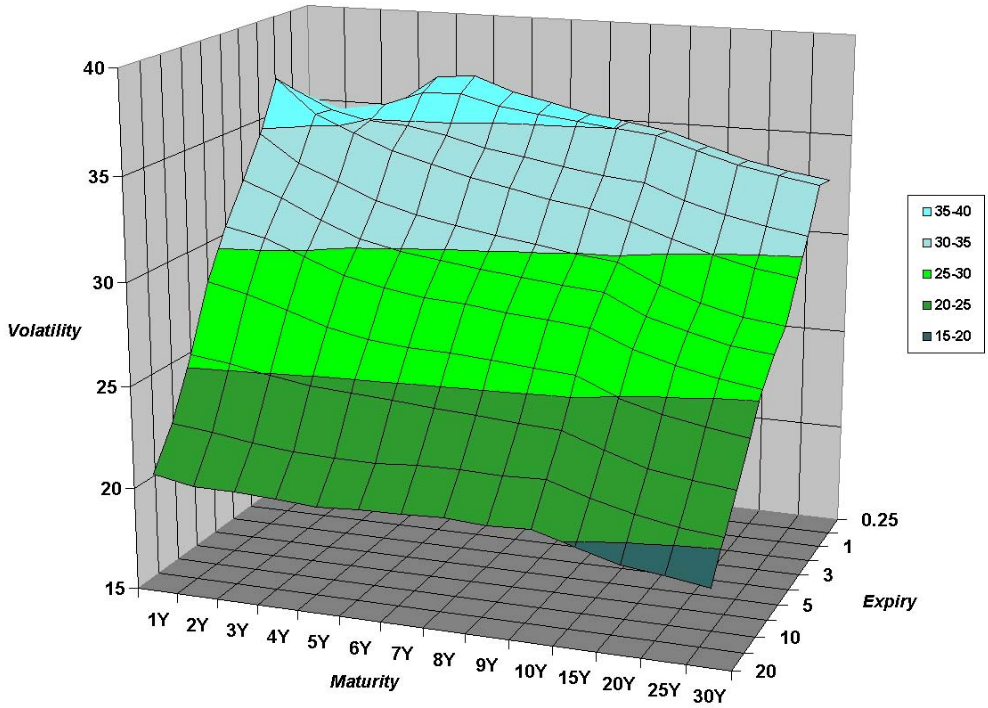

Figure 2 *Displaying swaption volatility as a function of option expiry and maturity using the surface plot in Microsoft® Excel. Again, all aspects of the plot can be edited interactively. Note, however, the irregular spacing between successive values on the axes, which is not reflected in the physical distance between them.*

For a contour plot, lines are drawn which connect together all places between the axes where the dependent variable has a certain value (sometimes called the *threshold*); different thresholds generate different lines, and their shape indicates something about the variation in the dependent variable. A surface plot uses a third dimension, at right angles to the dimensions used for the axes; the height of the surface at any point is proportional to the value that the dependent variable has at that point. These techniques are available in Excel.

Figure 2 shows a plot of volatility as a function of maturity and expiry using the surface chart type in Excel. As in the case of the plot shown in Figure 1, labels and colors can be interactively edited to produce the desired display (here, contours have also been added to the surface by coloring it using various threshold values for the dependent variable; this can be helpful for quantitatively determining regions of interest). Other options for display are also possible for this chart type. This is a three-dimensional object, and so some control over the viewpoint is required. Excel allows this in a somewhat indirect fashion via the 3-D View dialog box, but a more interactive interface (e.g. by clicking and dragging with the mouse) would perhaps be more helpful in some circumstances.

There is a problem with the display in Figure 2. To see it, consider this snapshot of the original dataset (as taken from, for example, a spreadsheet within Excel – here, maturity and expiry are both in years [Y]):

Volatility as a function of underlying maturity (top) and option expiry time (left)

	1Y	**2Y**	**3Y**	**…**	**10Y**	**15Y**
0.25Y	29.96	34.06	35.20	…	34.06	35.20
0.5Y	32.93	34.90	34.50	…	34.75	34.14
1Y	36.28	35.38	34.93	…	32.81	32.07
…	…	…	…	…	…	…
10Y	25.55	25.27	24.92	…	23.76	23.05
15Y	22.73	22.50	22.19	…	22.07	21.45

As can be seen from the table, the spacing between successive values for the independent variables is irregular – e.g., the gap between the first and second expiries is 0.25, but that between the second and third is 0.5. One reason for this might be that it had been deemed that the data should be closer together for low values of maturity and expiry time because the volatility changes faster in that region. Or it might be because not as much detail was required about the behavior of the volatility at long times and maturities. The point of interest, however, is that these irregularities are not truly reflected in the Excel plot in Figure 2. Although the labels on the axes correctly correspond to the values of the independent variables for each row or column of the table, the physical spacing between them (which is everywhere the same) does not. Thus, for example, the gap between the first and second label on the expiry axis corresponds to a change in value of 0.75, but that between the second and third represents a change of 2. This can be misleading, since the physical spacing between is the same for both intervals.

Of course, by paying attention to the shape of the surface and the labels on the axes, the viewer can overcome this problem (for example, in Figure 2, the somewhat abrupt change in slope of the front edge of the surface at the 10Y maturity mark could be attributed to the distortion caused by the five-fold change in spacing which takes place here). Care must be taken however, to ensure that a correct interpretation of the shape is made and it would be perhaps be better to visualize datasets such as these without artifacts like this, since they could mask interesting behavior in the dataset itself – which is, after all, the thing that we're trying to find.

A shortcoming in Excel's surface display, which is closely related to this, is its inability to display discontinuities in the surface. These arise when the dependent variable can take on two different values for the same value of a dependent variable (or – sometimes in numerical practice – two values of the dependent variable that are extremely close together). An example occurs in solutions of the Black-Scholes equation that model the value of an option as a function of stock price and time to maturity; at certain maturities, a (discrete) dividend is paid causing the value of the option to fall. This shows up as a vertical step in the surface – but not in Excel, which makes no allowance for changes in the value of the spacing between independent variables.

As if this wasn't bad enough, further complications can be introduced to our example. The volatility actually depends on the strike price as well as expiry and maturity – i.e. it is a function of three independent variables. We could produce a series of plots like Figure 2 for other values of the strike, but analysts are typically interested in other relationships as well – for example, the plot of volatility as a function of strike at fixed expiry and maturity, which sometimes includes the minimum known as the "volatility smile".

In the second article in this series, we'll take a look at how some of the problems that we've found in the display of financial data can be addressed through the use of some of the more advanced visualization techniques that are available in a dedicated visualization software package.

Dr Jeremy Walton (jeremy.walton@nag.co.uk) is a Senior Visualization Consultant with the Numerical Algorithms Group (NAG, www.nag.com), a worldwide organization dedicated to developing quality mathematical, statistical and 3D visualization software components that are used by most major finance houses in the world.

From the March/April 2004 issue of Financial Engineering News.