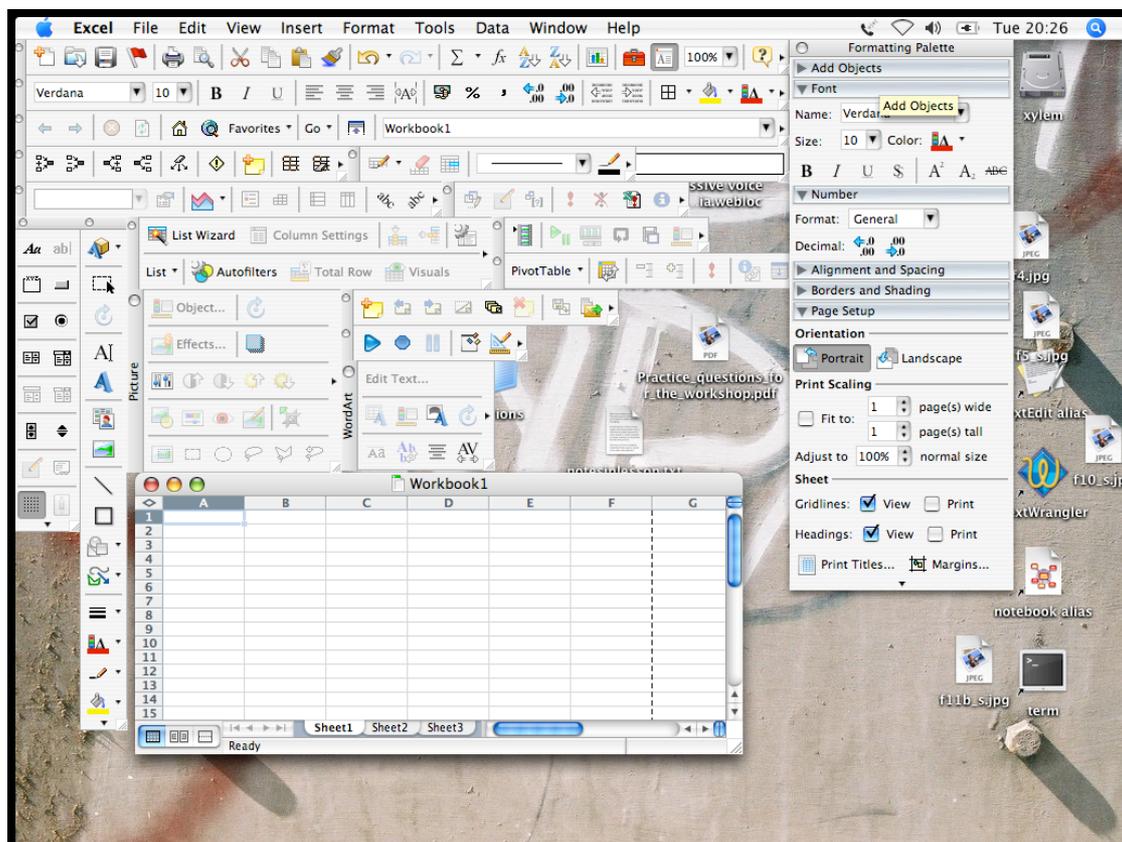


## Plotting scientific data in MS Excel 2003/2004



The screen grab above shows MS Excel with all the toolbars switched on - remember that some options only become visible when others are activated. We only use about 5% of the functions in Excel. The problem is, everyone uses a *different* 5%.

This handout explains plotting scientific data using Excel. Marketing managers, accountants, and engineers will have different requirements and will use different functions.

My screen shots were all obtained using MS Excel 2004 for the Mac OS platform, but the menus and functions are identical in MS Excel 2003 for Windows XP.

If you do not have MS Office, you can download OpenOffice for nothing from <http://openoffice.org/>

The Calc spreadsheet component can be configured to work in a similar way to Excel 2003, and can open and save files in Excel format.

## **The data sets**

To plot graphs you need to have some data to plot. The two data sets below were quoted in Lucy, 2005, and these data sets are from a Forensic Science context.

### **Data Set 1**

Time since discharge in days of a shotgun cartridge and height of a peak associated with nitroglycerin. The concentration of nitroglycerin decreases with time, and this can help estimate the time since discharge.

Time since discharge / days	Nitroglycerin / arbitrary peak height unit
1.21	218.34
2.43	216.16
3.62	100.00
4.69	75.55
7.49	56.52
9.42	50.62
11.60	31.00
14.69	41.44
21.50	15.53
25.70	14.63
29.86	10.41
37.20	5.16
42.42	7.26

### **Data Set 2**

Methyl Violet is a common dye used in inks dating from the 1950s. Documents were artificially 'aged' using strong UV light for various periods of minutes. The resulting average molecular weight ( in Daltons) of the dye was measured using laser desorption mass spectroscopy.

UV exposure time / min	Average molecular weight / Da
0.0	367.20
15.3	368.97
30.6	367.42
45.3	366.19
60.2	365.91
75.5	365.68
90.6	365.12
105.7	363.59

Turn over the page for instructions on how to plot data set 2 in MS Excel....

We need to

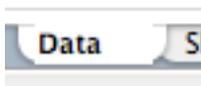
- **Type the data** into MS Excel
- Highlight the data columns and use the Chart Wizard, selecting the XY (Scatter) plot type
- Having completed the Chart Wizard, you will want to alter the colours used and the scales on the axes and add extra tick marks to allow readings to be taken from the graphs

### **Getting data into MS Excel**

- **Load MS Excel** and you should see a blank spread sheet open, showing a series of 'cells'
- Notice how each row has a number and each column a letter. A1 is the 'address' of the top left 'cell'
- Each cell can accept only one kind of data; either some text, *or* a number, *or* a formula. If you mix numbers and text, then Excel treats the combination as text.
- Click in the appropriate cells to add text or numbers
- Make your file look like the screen grab below - notice how the title text flows out of the cell A1 - but that text is still in the cell A1, and is not in B1 or C1

1	Average molecular weight of Methyl Dye for different UV exposures			D
2				
3				
4				
5				
6	UV exp time / min	Average mol wt / Da		
7	0	367.2		
8	15.3	368.97		
9	30.6	367.42		
10	45.3	366.19		
11	60.2	365.91		
12	75.5	365.68		
13	90.6	365.12		
14	105.7	363.59		
15				

- Type **enter** after each number and Excel drops down to the next row automatically
- Use the mouse or arrow keys to move across to the next column
- Note how **numbers align to the right** and **text to the left**
- Save your spreadsheet file to a floppy disc or to your USB stick drive
- See the 'tabs' at the bottom of the spreadsheet window - you can change the name of a tab - just click on it and type a new name



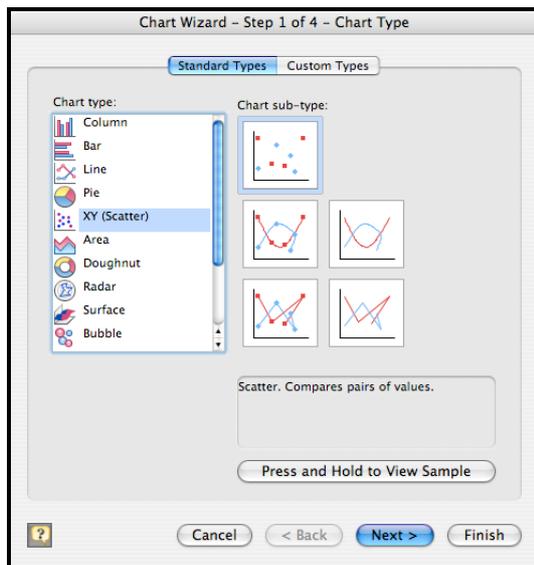
- Perhaps Data Set 2 would be a better name?

## Using the Chart Wizard to produce an XY (Scatter) plot

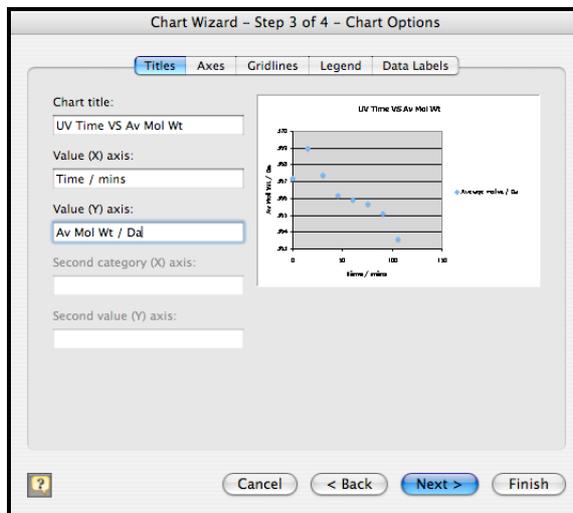
Data set 2 has unequal intervals between the X or *ordinate* data in column A. We must use the XY (Scatter) plot type in Excel. The other kinds of plot will produce an incorrect graph.

To start the chart wizard, you

- **Highlight** from cell A6 to B14
- Click on the **Chart Wizard toolbar button**  in the toolbar
- You should see something like this appear...

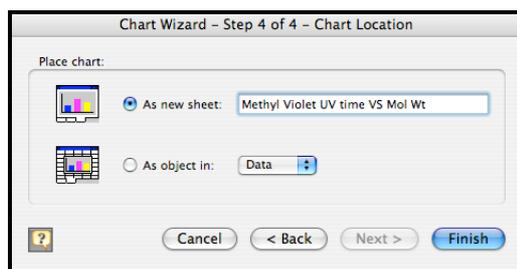


- Select the **XY (Scatter)** chart type, and pick the first Chart sub-type
- Click **Next** to continue the chart wizard...
- Click **Next** again as the **Data Range** is correct for this chart type
- Fill in the **title** and the X and Y axis **labels** as below...



- Click the **Legend** tab and untick the **Show Legend** tick box

- Click **Next** to advance to the final stage in the Chart Wizard



- I always tick **As New Sheet** and I give the chart sheet a sensible name
- This means that my chart is a whole page in the Excel workbook, so that my data takes as much space as possible on an A4 sheet
- Click **Finish** to insert your chart into your Workbook, and then Save your spreadsheet!

### **Colours and gridlines**

MS Excel uses a muddy grey background (known as the **Plot Area** in Excel), and often picks daft colours for the data markers. The gridlines also clutter up the graph to my thinking. You can change the colours and so on easily.

To clear the Gridline, just

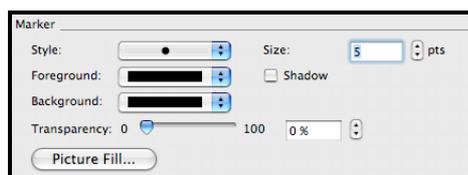
- Click on a **gridline** so the gridlines are selected
- Right click and select '**clear**' from the menu that appears

To change the background colour to white just

- Click on the muddy grey **Plot Area**
- Right click and select **Format Plot Area** from the menu
- Set the **Fill** colour to white and click **OK**

To change the colour of the data points themselves, just

- **Click** on any data point once
- All the data points are highlighted
- **Right click** and select **Format Data Series** from the menu
- The Colors and Lines tab should be open
- Look at the **Marker** color of the dialogue box



- Untick the **Shadow** check box - shadow effects and picture fills have no role in a scientific graph
- Pick a dark fore and background colour and click OK to change the graph

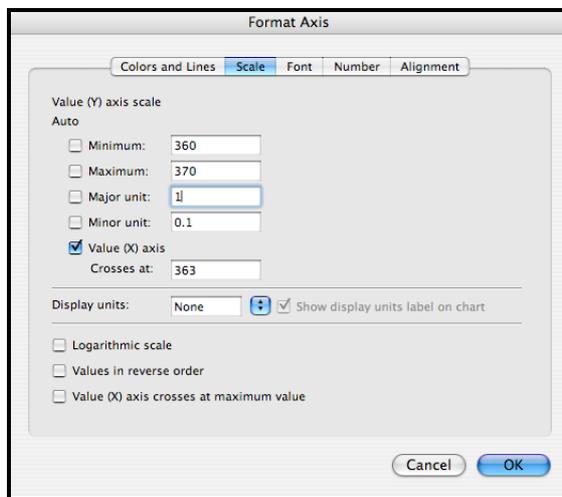
## Scales and markers

Excel makes a reasonable guess at the scales for data set 2, but sometimes, Excel uses far too small a scale.

You can add 'tick marks' to the axes in between the numbers to make it easier to read results from a graph.

To change the Y axis scale to start from 360, you just

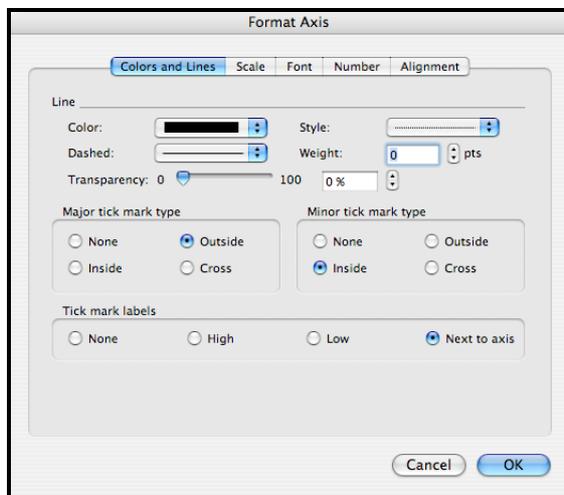
- **Double click** on the Y axis
- The **Format Axis** dialogue box should appear
- Click on the Scale tab



- Untick the **Auto** setting for **Minimum, Maximum, Major Unit** and **Minor unit**
- Set the values as shown in the screen grab above: setting the Minor unit to 0.1 will give 10 divisions between the numbers on the axis, just like a paper graph
- Click OK

To show the tickmarks...

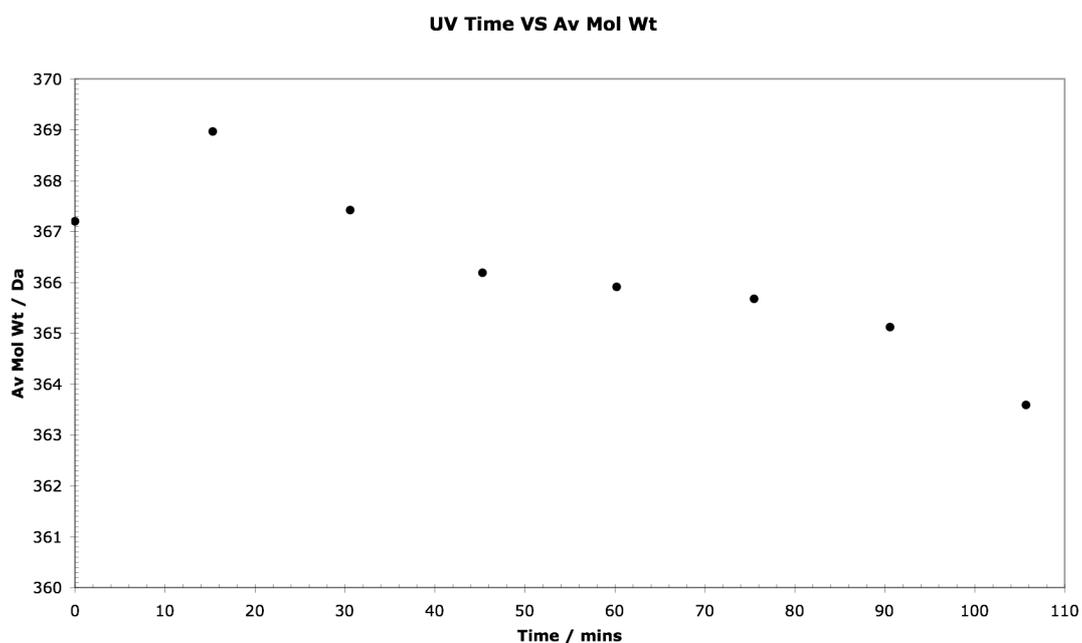
- Click on the **Colors and Lines** tab
- You should see something like this...



- Notice how I have ticked the Minor Tick mark type to give 'inside ticks', small marks on the Y axis.

## Your turn

Alter the X axis so that the scale starts from 0 and ends at 110, with major tick marks every 10 units and minor tick marks every 2 units.



My graph ended up looking like the one above... but the illustration has been reduced in scale.

## **Homework**

- Task 1: Make a new table and chart for data set 1. How would you describe the shape of the chart?
- Task 2: Experiment with colours for the data points themselves. There is a combination of foreground and background colour that will give crosses and open circles. Find the combination!\
- Task 3: Experiment with the scale dialogue. Try silly combinations of scale and see what happens.
- Task 4: Use MS Excel help to learn about 'Trend lines'. Add a linear 'trend line' to your plot for Data set 1. The usual name for trend line is Line of Best Fit.

## **References**

Lucy, David, *Introduction to Statistics for Forensic Scientists*, John Wiley and Sons, Chichester, 2005