RGB EXTRACTION AND DATA ANALYSIS

The software used to extract the RGB values called **ImageJ**. It is completely free software which is widely used in the field of microscopy (for further information please refer to 'Digital Image Processing' by Burger and Burge).

The download page is here:

http://rsb.info.nih.gov/ij/download.html

Which download depends on your operating system and whether you have Java installed or not. The rest of this procedure assumes that the software is now installed and ready to use. This procedure was written using Windows 7 32-bit, the software is platform independent so the procedure should be the same across platforms. This procedure also uses Microsoft Excel 2010 and a basic image editing software (such as Irfanview).

STEP 1. – Import Images from Scanner

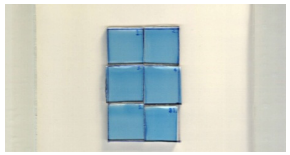
Using the USB to mini-USB cable, import the pictures from the hand-held scanner. To do this plug it into the PC and turn on the scanner. The scanner should be recognised by the PC and allow access to the picture folder. Copy and Paste the pictures to a new folder on the desktop or wherever is convenient on your machine.

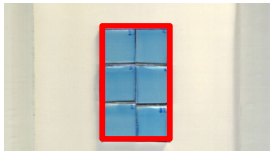
STEP 2. – Cropping the pictures

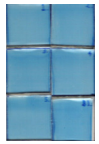
Make a new folder within the folder of imported pictures.

Using basic image editing software crop all of the pictures to the outer edges of the samples.

Like so:

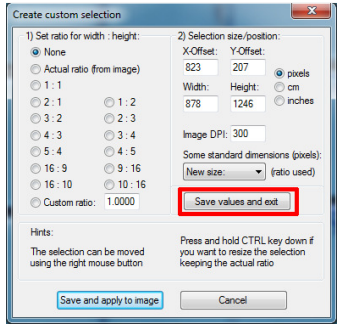






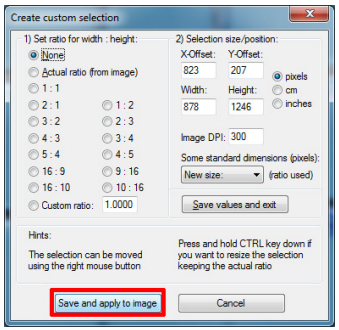
This is achieved relatively easily and quickly using Irfanview. Note that some of these steps have shortcuts associated with them and learning to use them will make this part of the analysis much quicker.

1. Open the software and ensure that **View > Display Options (Window Mode) > Fit only big images to window** is selected.
2. Open the first image in the sequence with Irfanview and draw a selection around the outer edge of the samples. Now select **Edit > Custom crop selection…**. The following window will pop up:

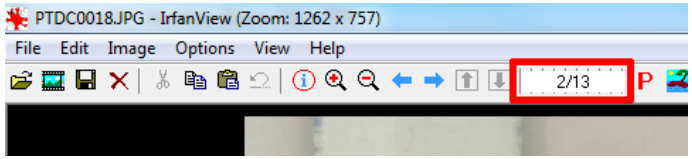


Select the **Save values and exit** button (marked in red above). What this does is save the selection dimensions so that for every picture you can recall this custom selection and apply it to each picture in turn – like a cookie cutter. This ensures that every picture is the same size and this will help in analysis.

3. Once this is done, using the **right mouse button** you can move the selection to include all the samples (if it already doesn't). If you accidentally left-click the selection, it will just zoom to the selection. If this happens, press the letter '**F**' (F for fit) to refit the image in the window and then press **Shift + C** to apply the custom selection again.



4. Now select **Edit > Crop selection**. You will notice that the picture outside of the samples has disappeared leaving only the samples. Click **File > Save as…** (shortcut letter '**S**') and save the picture in its own folder.
5. It is prudent to name the picture the irradiation time in seconds which it represents.
6. Once the cropped picture is saved, press the right arrow key on the keyboard to skip to the next picture in the series. The number in the series is displayed in the toolbar at the top of the window:



Steps 3 to 6 can now be repeated quickly with shortcuts (provided steps 1 and 2 have been done):

Open Irfanview,

press '**O**'

browse to and select the desired picture

'**Shift + C**' and move selection (with right mouse button!) so samples fit inside.

'**Cntrl + Y**' to crop to selection.

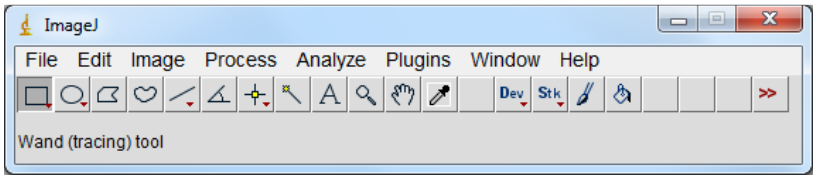
Press '**S**' key to save (name it the time in seconds)

Press right arrow key to move to next image.

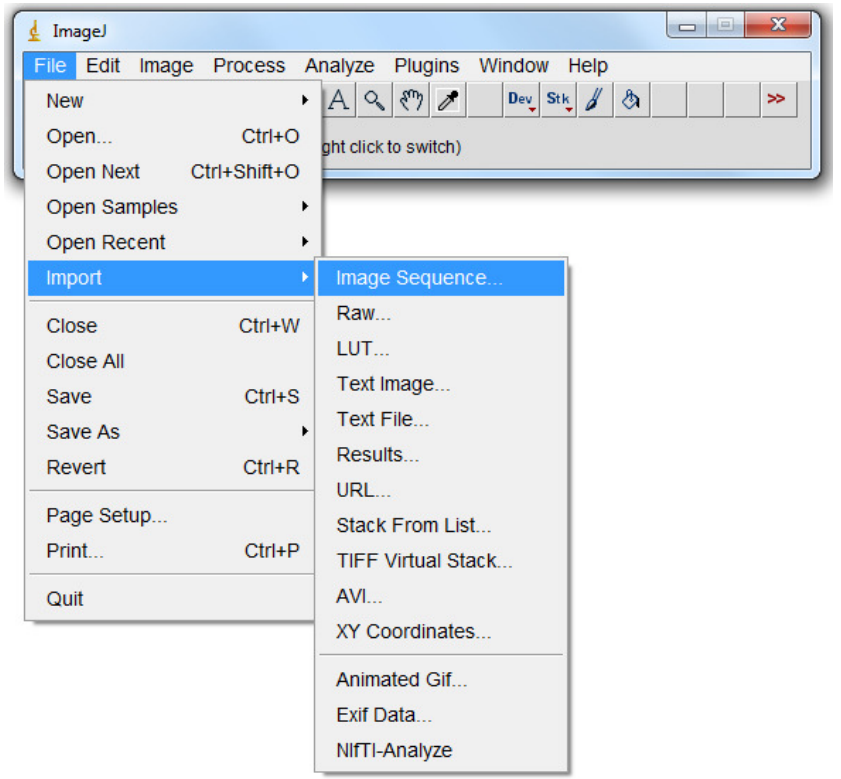
NOTE: Be careful to save each cropped picture in the correct order and **to its own folder**. Do this for all the pictures, they should look like this:

STEP 3. – Image analysis with ImageJ Software

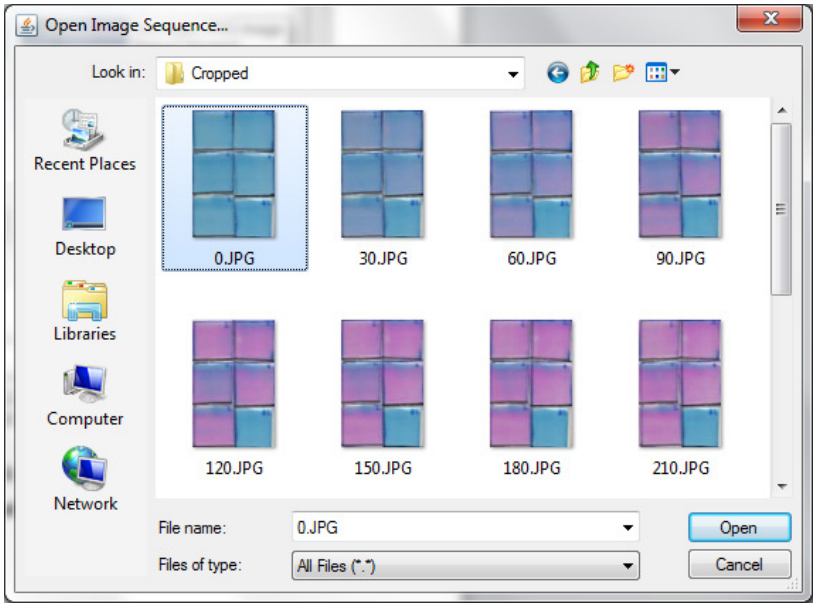
1. Open ImageJ:



2. Click '**File > Import > Image Sequence…**' like so:

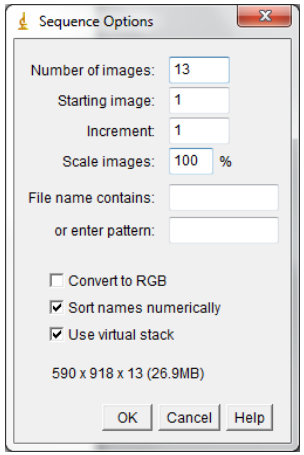


3. An explorer dialogue box will open, browse to the folder of cropped images:

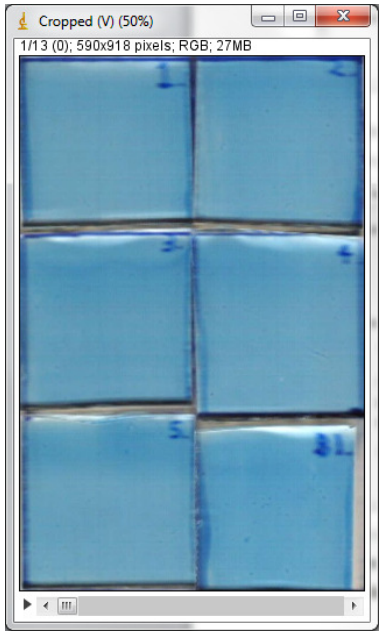


Select the first picture in the series only and then click '**Open**'.

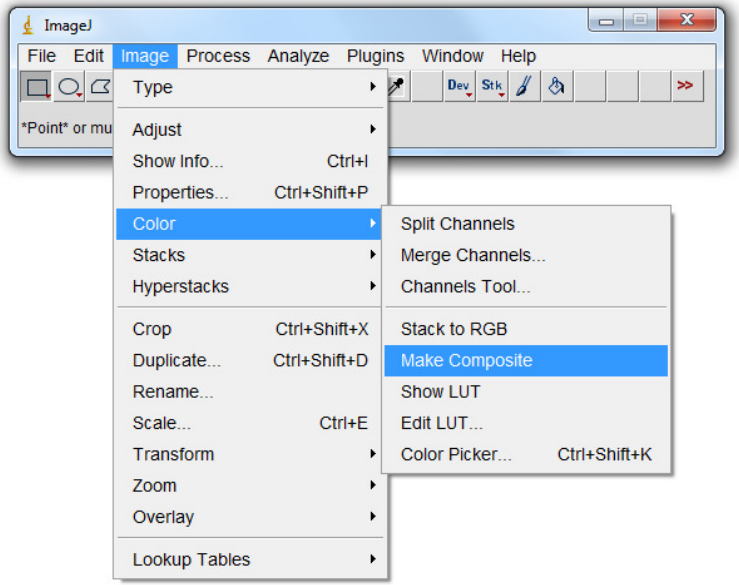
The following box will pop-up, select the '**Use virtual stack**' check box and click '**OK**':



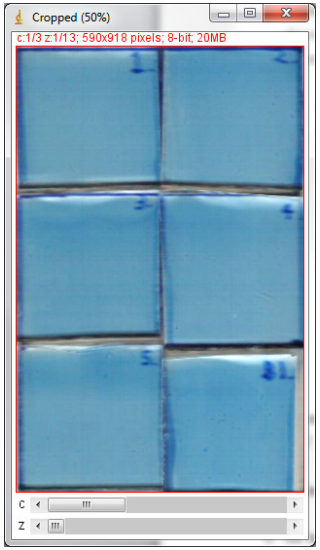
4. This imports the entire series of pictures as a stack. You can see each image in the series by clicking the scroll bar at the bottom of the window, when you are finished checking the images make sure the scrollbar is set all the way to the left:



5. Go back to the main ImageJ window and select '**Image > Color > Make Composite**'.



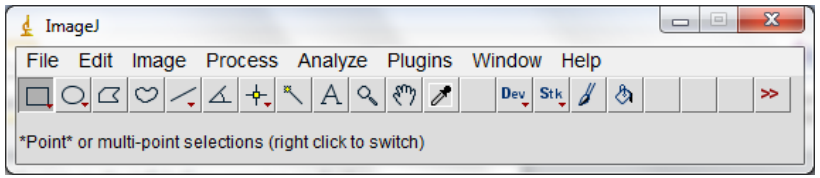
ImageJ splits each picture into its Red, Green and Blue channels. A pop-up will appear to this effect, ignore or close it. You will notice that the image window (4., above) will now have a red outline and a second scroll bar labelled 'C' appears at the bottom like so:



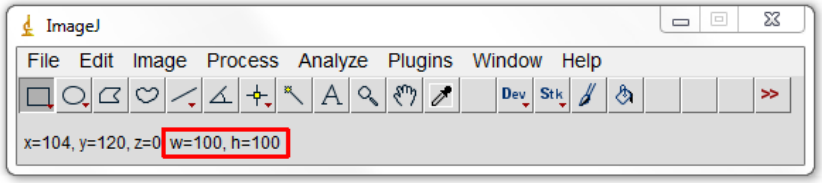
This red outline indicates which channel is shown and the '**C**' scroll bar indicates the 'component' channels which can be scrolled through for each image – although this is not necessary.

NOTE: Make sure that both scroll bars are all the way to the left before carrying on further.

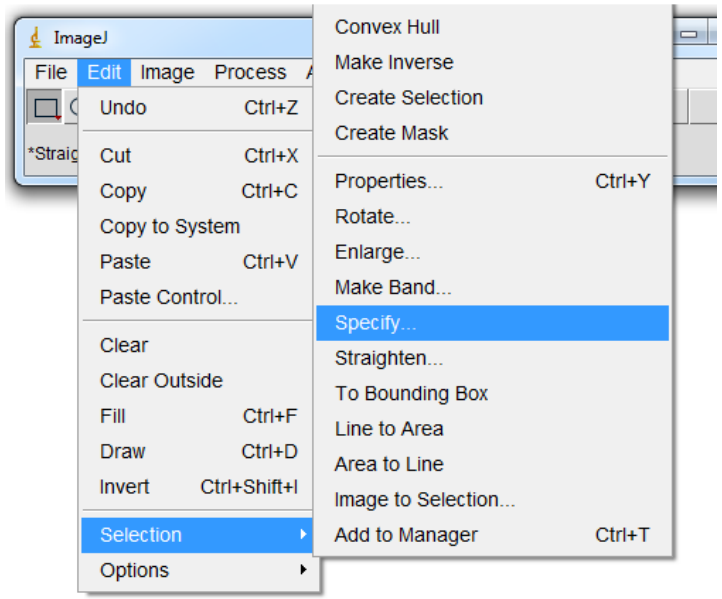
6. Go back to the main ImageJ window and select the Rectangular selection button on the main toolbar (it should already be selected).



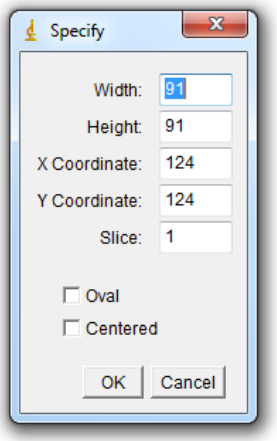
Hold the Shift key down and select an area of 100 x 100 pixels in the centre of the first glass sample in the picture, to help with this you will notice in the main ImageJ window it tells you what size the selection is, like so:



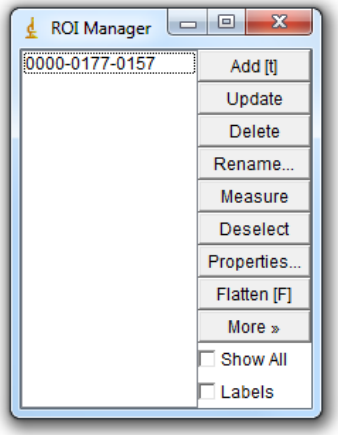
This may be difficult to do if the mouse is not sensitive enough in which case go to the main window and select '**Edit > Selection > Specify…**' like so:



The following window will pop-up, make the width and height each 100 and click '**OK**'.



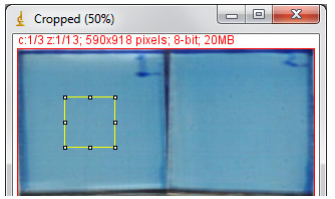
7. You can move the yellow square in the picture to wherever is needed, however if you click outside the selection (and on the picture) it will disappear and it will have to be drawn again. Note that it is not necessary to grab the actual yellow box, just click inside the box to be able to move it. Scroll through the pictures using the 'Z' scroll bar at the bottom of the picture window, to be sure that the selected area is over the sample and the pictures aren't moving around too much (not catching an edge of the sample or the blue label for instance).
8. To save the selection as a 'Region of Interest (ROI)' press '**Ctrl + T**' and the ROI Manager will pop-up:

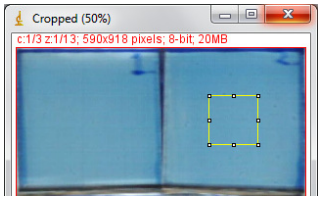


In the column on the left there is a number which represents the region of interest (yellow selection box) in the picture. Click '**Rename…**' and name it to whatever glass sample the selection is on.

Now, go back to the picture window and move the selection to the next sample as shown below.

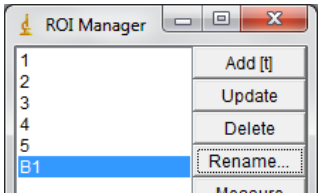
Before: After:

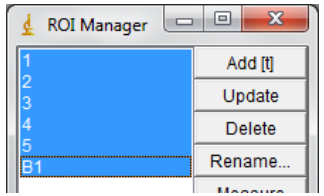




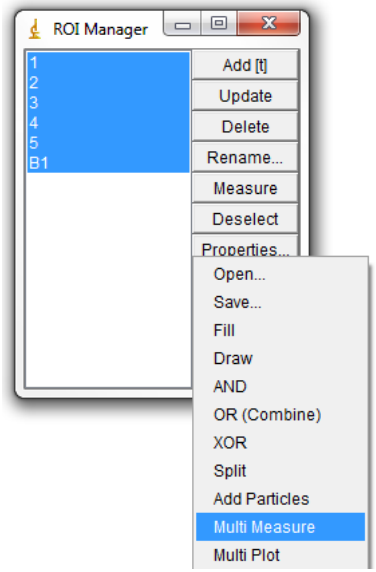
Press '**Ctrl + T**' again. Once again a number will appear in the ROI Manager window, rename it (in this case to '**2**'). Do this for all the samples in the picture.

9. There should now be 6 ROIs listed in the ROI Manager window (one for each sample, including the blank B1). Now the RGB data for each of those areas in each of the pictures can be extracted and saved. Select the first ROI in the window, hold the shift key down and click the last one. The entire list of ROIs will now be highlighted.

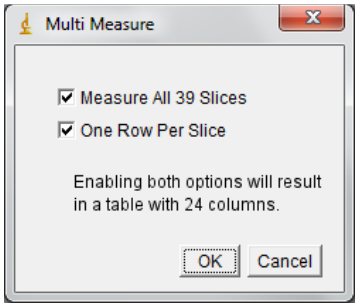




Now click **More… > Multi Measure**'.

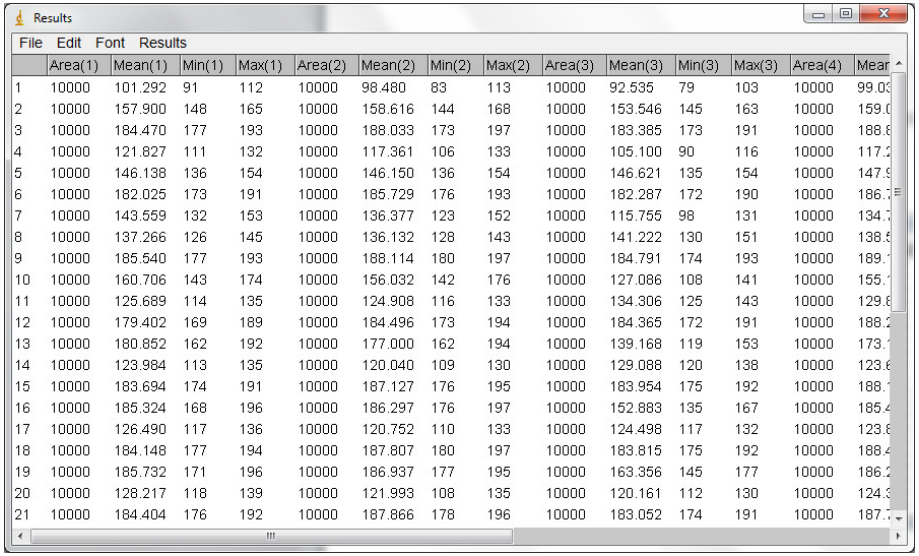


The following window will pop-up:



Click '**OK**'

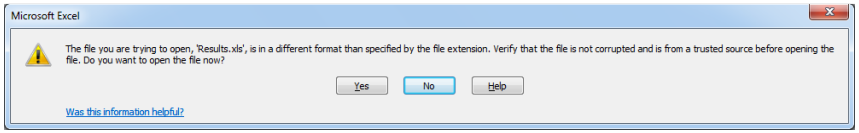
The following Results window will pop-up after clicking OK. In this window click '**File > Save As…**' or '**Cntrl + S**' and save the data in a folder. Note that the data are saved as a '.xls' file (Office 2003) as opposed to a '.xlsx' file (Office 2007/2010).



	Area(1)	Mean(1)	Min(1)	Max(1)	Area(2)	Mean(2)	Min(2)	Max(2)	Area(3)	Mean(3)	Min(3)	Max(3)	Area(4)	Mean
1	10000	101.292	91	112	10000	98.480	83	113	10000	92.535	79	103	10000	99.0
2	10000	157.900	148	165	10000	158.616	144	168	10000	153.546	145	163	10000	159.0
3	10000	184.470	177	193	10000	188.033	173	197	10000	183.385	173	191	10000	188.8
4	10000	121.827	111	132	10000	117.361	106	133	10000	105.100	90	116	10000	117.2
5	10000	146.138	136	154	10000	146.150	136	154	10000	146.621	135	154	10000	147.9
6	10000	182.025	173	191	10000	185.729	176	193	10000	182.287	172	190	10000	186.7
7	10000	143.559	132	153	10000	136.377	123	152	10000	115.755	98	131	10000	134.7
8	10000	137.266	126	145	10000	136.132	128	143	10000	141.222	130	151	10000	138.5
9	10000	185.540	177	193	10000	188.114	180	197	10000	184.791	174	193	10000	189.1
10	10000	160.706	143	174	10000	156.032	142	176	10000	127.086	108	141	10000	155.1
11	10000	125.689	114	135	10000	124.908	116	133	10000	134.306	125	143	10000	129.8
12	10000	179.402	169	189	10000	184.496	173	194	10000	184.365	172	191	10000	188.2
13	10000	180.852	162	192	10000	177.000	162	194	10000	139.168	119	153	10000	173.1
14	10000	123.984	113	135	10000	120.040	109	130	10000	129.088	120	138	10000	123.6
15	10000	183.694	174	191	10000	187.127	176	195	10000	183.954	175	192	10000	188.1
16	10000	185.324	168	196	10000	186.297	176	197	10000	152.883	135	167	10000	185.4
17	10000	126.490	117	136	10000	120.752	110	133	10000	124.498	117	132	10000	123.8
18	10000	184.148	177	194	10000	187.807	180	197	10000	183.815	175	192	10000	188.4
19	10000	185.732	171	196	10000	186.937	177	195	10000	163.356	145	177	10000	186.2
20	10000	128.217	118	139	10000	121.993	108	135	10000	120.161	112	130	10000	124.3
21	10000	184.404	176	192	10000	187.866	178	196	10000	183.052	174	191	10000	187.7

STEP 4. – Data Processing

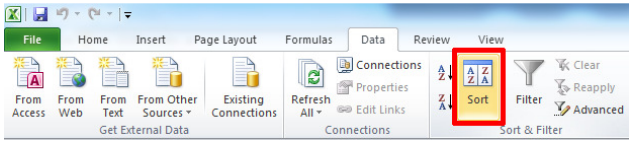
There should now be a "Results.xls" containing the mean RGB values from each of the pictures in the series. Note that the data file is an older format excel file (.xls) so Office 2007/2010 users will have a pop-up to deal with:



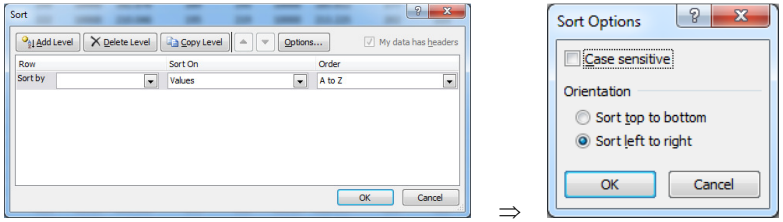
Click '**Yes**'.

Press '**Ctrl + A**' to select all the values in the sheet.

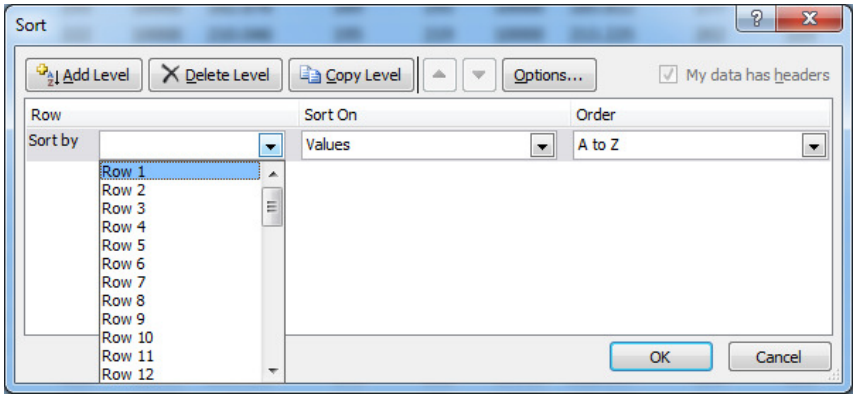
Click the '**Data**' Tab and select the '**Sort**' button.



In the pop up click the '**Options…**' button and select the '**Sort left to right**' radio button and click '**OK**':

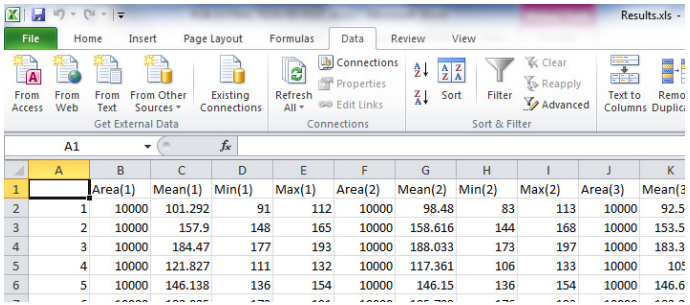


Now select '**Row 1**' from the '**Sort by**' drop down menu (in the 'Sort' window) and click '**OK**':

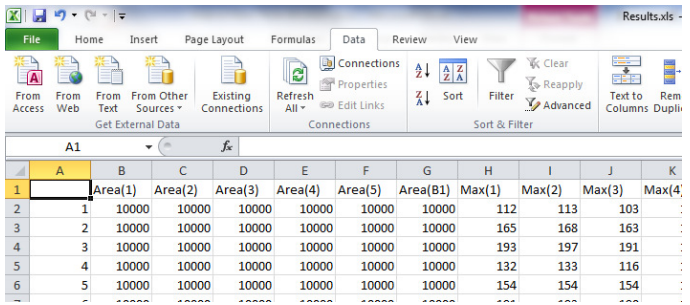


The data will now be sorted, left to right:

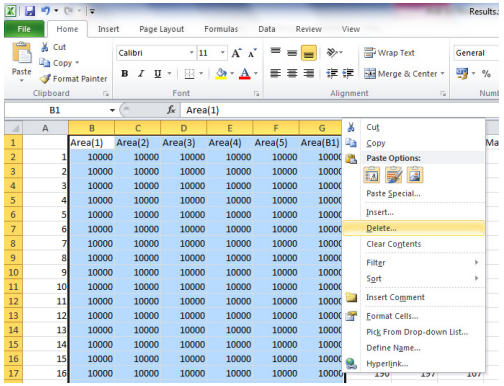
Before sort:



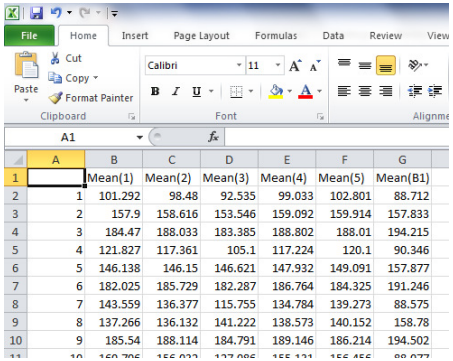
After sort:



The columns of data of interest are the ones titled "**Mean(xx)**" etc.. all the rest can be deleted. Multiple columns can be selected at a time by clicking on the grey column title, the cursor should change to a downward pointing black arrow, once selected right-click the columns and select '**Delete**'.



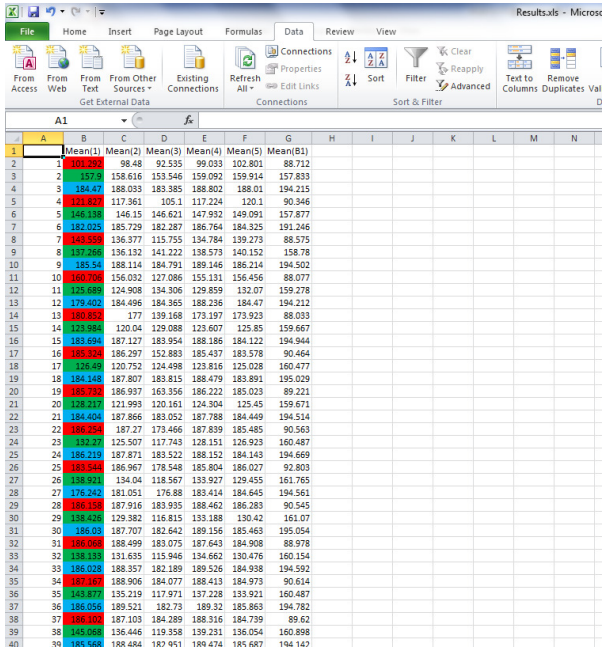
The only columns of data that should be left now are the ones titled '**Mean(xx)**' (where xx is the number of the sample) and a single column to the left (column A – if this was deleted right-click column A and click '**Insert**').



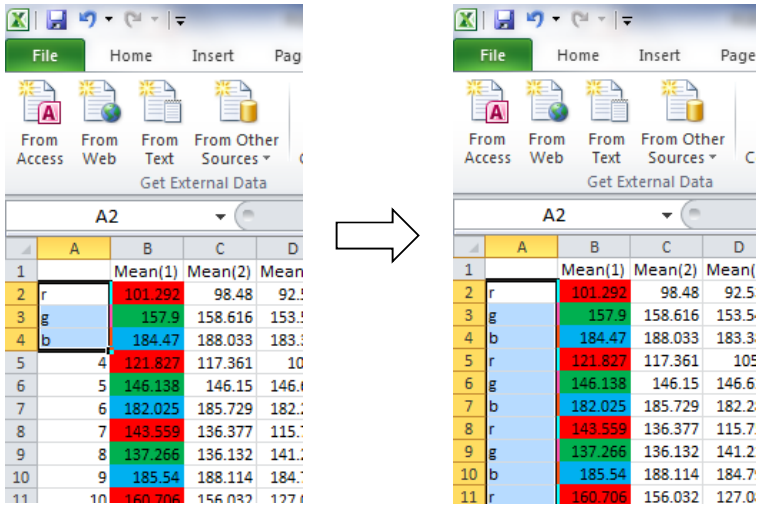
	A	B	C	D	E	F	G
1		Mean(1)	Mean(2)	Mean(3)	Mean(4)	Mean(5)	Mean(B1)
2	1	101.292	98.48	92.535	99.033	102.801	88.712
3	2	157.9	158.616	153.546	159.092	159.914	157.833
4	3	184.47	188.033	183.385	188.802	188.01	194.215
5	4	121.827	117.361	105.1	117.224	120.1	90.346
6	5	146.138	146.15	146.621	147.932	149.091	157.877
7	6	182.025	185.729	182.287	186.764	184.325	191.246
8	7	143.559	136.377	115.755	134.784	139.273	88.575
9	8	137.266	136.132	141.222	138.573	140.152	158.78
10	9	185.54	188.114	184.791	189.146	186.214	194.502

The data have to be sorted once more.

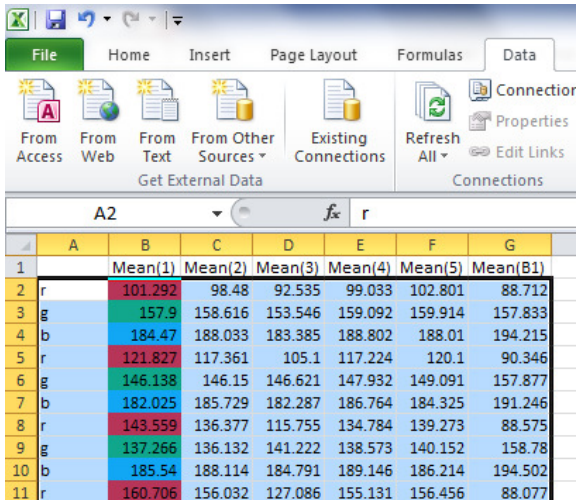
The data in each column are present as a list of the Red, Green and Blue (RGB) values of each of the pictures just taken – these have been coloured in the picture below to illustrate this. The values can be sorted into the RGB values by typing in '**r**' into cell A2, '**g**' into cell A3 and '**b**' into cell A4.



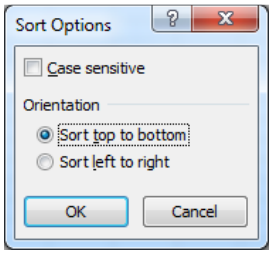
Select these cells and copy them down so that you get r,g b repeating down column A like so:



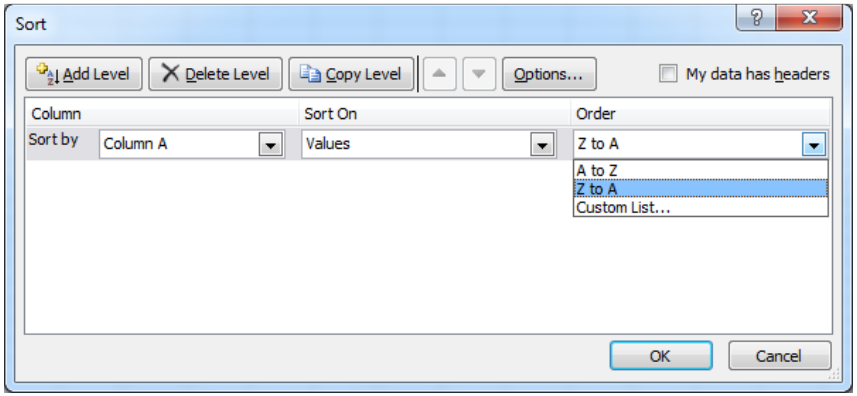
Now select the data set (not including the titles) as below:



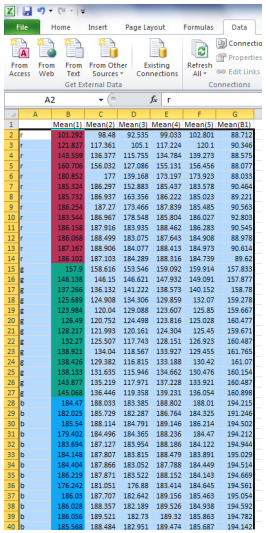
Go to the 'Data' tab and click the '**Sort**' button. Now click the '**Options…**' button and click the '**Sort top to bottom**' radio button and click '**OK**'.



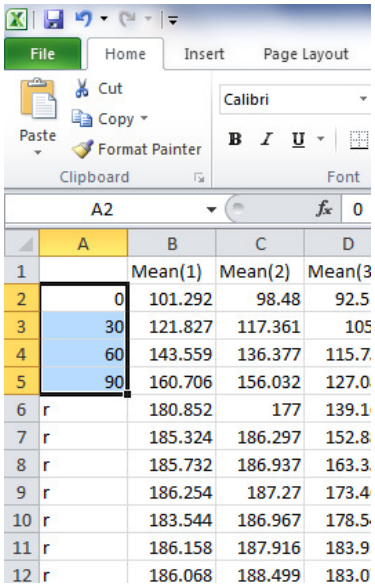
Now, in the '**Sort by**' drop down menu select '**Column A**' and from the '**Order**' drop down menu select '**Z to A**'.



The values should now look like this:



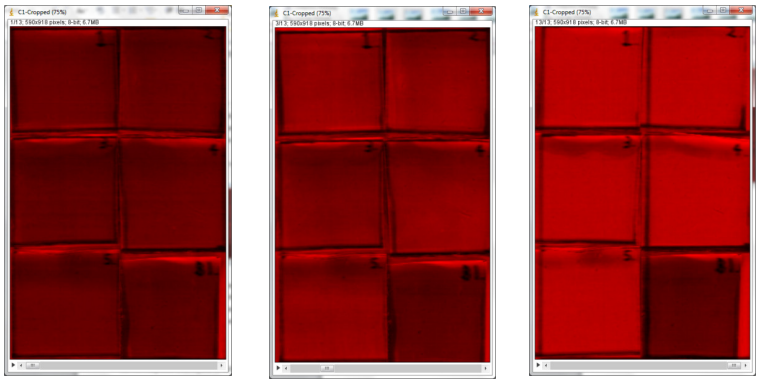
The pictures were taken every 30 seconds, so column A was filled with the time in seconds like so:



	A	B	C	D
1		Mean(1)	Mean(2)	Mean(3
2	0	101.292	98.48	92.5
3	30	121.827	117.361	105
4	60	143.559	136.377	115.7
5	90	160.706	156.032	127.0
6	r	180.852	177	139.1
7	r	185.324	186.297	152.8
8	r	185.732	186.937	163.3
9	r	186.254	187.27	173.4
10	r	183.544	186.967	178.5
11	r	186.158	187.916	183.9
12	r	186.068	188.499	183.0

The red values are the ones of interest as this is the region of the spectrum which Resazurin absorbs most strongly (λ_{max} = 610 nm). The fact that the colours are additive means that as there is less and less Resazurin to absorb the red light and there is a consequent increase in the red component of the pictures of the samples as they are irradiated.

This can be seen in a picture of the red component only for increasingly irradiated samples:



The blanks in the bottom right hand corner do not reduce the dye from blue to pink and hence the red component does not increase with irradiance.

The maximum value the red component can reach is 255 which represents the maximum of 256 possible values (0 to 255) contained in an 8-bit per channel digital image. The red values are rarely more than 200.

Typical response and Data analysis

The red component from a piece of self-cleaning glass is plotted vs. irradiation time. It can be seen that the values reach a maximum and then level off so that a clear break in the line is observed. Perhaps the most accurate way then to determine the time of the end point is to plot a line through the linear portions of the curve and determine the time at which they intersect.

A plot of a typical response is shown in **Figure 1.** below.

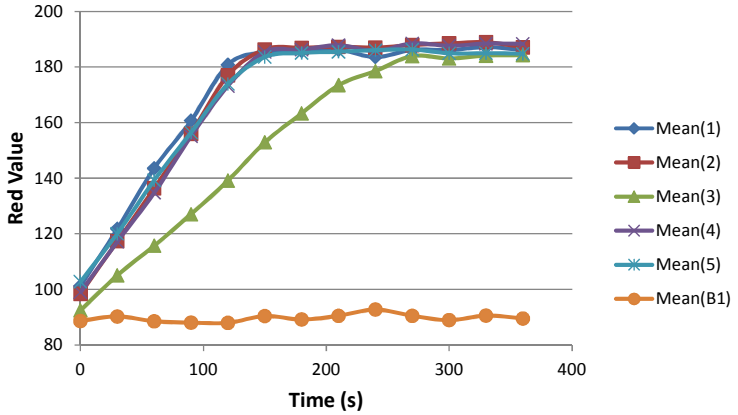


Figure 1(a). – The variation of red component as a function of irradiation time (s) for a typical self-cleaning glass sample.

Usually there is a clear break in the response curve (**Figure 1 (a)**) so that they can roughly be divided into two linear portions; before the end-point and after the end-point as shown in **Figure 1 (b)**.

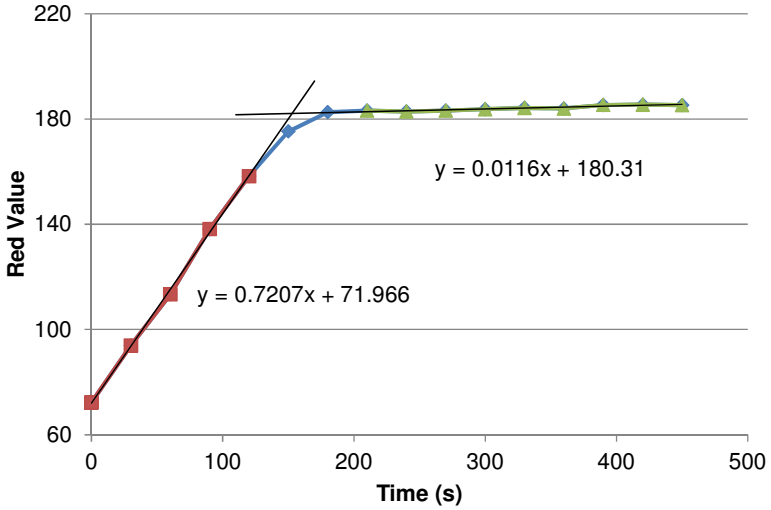


Figure 1(b). – The linear portions of the data plotted in different colours are the portions used to determine the end point.

Plotting a trendline through each part of the curve gives an equation for each trendline. So, for example using the equations shown in the plot above

$$Y = 0.7207x + 71.966 \textbf{ …eq. 1.}$$

$$Y = 0.0116x + 180.31 \textbf{ ...eq. 2.}$$

Where these two lines intersect, (x, y) will be the same so the equations equal each other:

$\Rightarrow 0.7207x + 71.966 = 0.0116x + 180.31$

Rearranging to find x:

$\Rightarrow 0.7207x - 0.0116x = 180.31 - 71.966$ …at this intersection, x is the same value so:

$\Rightarrow 108.344 = 0.7091x$

$\Rightarrow x =$ **<u>152.79 seconds</u>**

To test this, this value can be inserted into either of equations 1 or 2 above and then the value of *y* for the above calculated value of *x* can be obtained.

For example:

Y = 0.7207 × 152.79 + 71.966 = **182.08 (value of red component at the end point time of 152.79 seconds)**

The coordinates (152.79, 182.08) are shown plotted in **Figure 1 (c)** below. *The only value of importance here though is the time for the end point (x).*

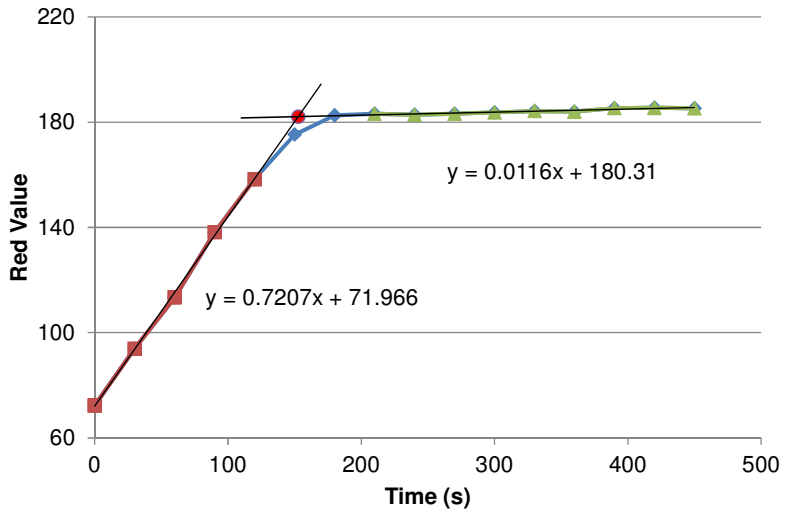


Figure 1(c). – The calculated intersection plotted on the data (the red dot).

This can be summarized as:

$X_i = (C_1 - C_2)/(M_2 - M_1) = \Delta C/\Delta M$

Where the equation is of the form y = mx + c.

Thankfully in excel this process can be done without having to plot trendlines for every set of data by using the SLOPE and INTERCEPT functions.

Draw the following table in excel:

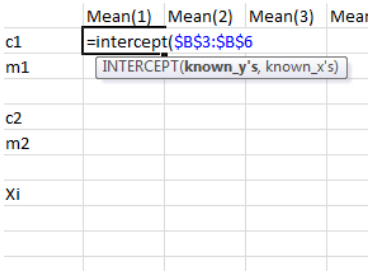
	Mean(1)	Mean(2)	Mean(3)	Mean(4)	Mean(5)	Mean(B1)
c1						
m1						
c2						
m2						
Xi						

where **c1** and **m1** represent the intercept and gradient for the first, linear part of the curve. Similarly, **c2** and **m2** represent these in the second, linear portion. Xi represents the value of x at the intersection of the two linear trendlines – i.e. the end point.

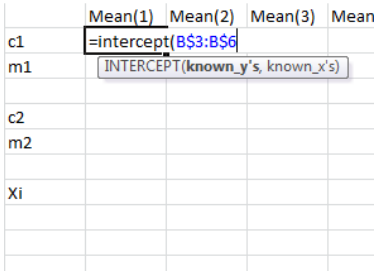
Type in "=intercept(" (without the inverted commas) and select the linear portion of the first part of the dataset, as shown in the picture above (the Y values first and the X values second, separated by a comma – depending on your country). So that the equation can be copied across, after selecting the y-data (shown in blue) press the F4 key twice. Type a comma ",". Select the x-data and press F4 once. Close the brackets and press 'Enter'.

Where the dollar sign "$" is in front of the column letter this fixes the column when copying. When the dollar sign is in front of the row number this fixes the row range, the difference is shown below.

Pressing F4 Once



Pressing F4 Twice



The equation should look like this when finished.

	Mean(1)	Mean(2)	Mean(3)	Mear
c1	=intercept(B$2:B$6,$A$2:$A$6)			
m1				
c2				
m2				
Xi				

Now copy across, hover the mouse over the bottom left corner (shown in red) and drag across for all the samples.

	Mean(1)	Mean(2)	Mean(3)	Mean(4)	Mean(5)	Mean(B1)
c1	102.0474					
m1						
c2						
m2						
Xi						

Like so:

	Mean(1)	Mean(2)	Mean(3)	Mean(4)	Mean(5)	Mean(B1)
c1	102.0474	97.9078	92.8784	98.6268	102.7906	89.474
m1						
c2						
m2						
Xi						

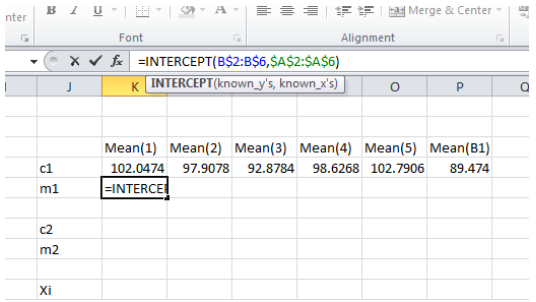
The intercept values should be roughly the same as the first points of the response curve and should be roughly comparable. If they are not it is likely that the equation has not copied properly because it is incorrect.

Now select the cell for the first sample in the series and copy it down into the row for **m1**:

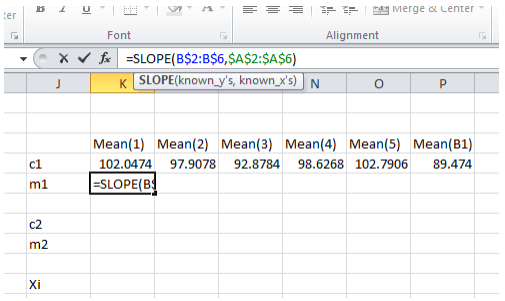
	Mean(1)	Mean(2)	Mean(3)	Mean(4)	Mean(5)	Mean(B1)
c1	102.0474	97.9078	92.8784	98.6268	102.7906	89.474
m1						

	Mean(1)	Mean(2)	Mean(3)	Mean(4)	Mean(5)	Mean(B1)
c1	102.0474	97.9078	92.8784	98.6268	102.7906	89.474
m1	102.0474					

Edit the equation by deleting the 'intercept' term and replacing it with 'slope':



Like so:



Copy the equation across all the samples as before and repeat for the second, linear portion of curve (usually the last points after the 'elbow' in the curve).

Now, for the last part, to determine Xi, the time for the end point is calculated by (C2 – C1) / (M1 – M2) like so:

	Mean(1)	Mean(2)
c1	102.0474	97.9078
m1	0.659997	0.65237
c2	185.3962	190.2396
m2	0.003103	-0.00677
Xi	=(K5-K2)/(K3-K6)	

Copy across for all samples as described previously.

	Mean(1)	Mean(2)	Mean(3)	Mean(4)	Mean(5)		
c1	102.0474	97.9078	92.8784	98.6268	102.7906		
m1	0.659997	0.65237	0.384173	0.620783	0.595333		
c2	185.3962	190.2396	181.6768	187.8599	190.0211		
m2	0.003103	-0.00677	0.00688	0.001107	-0.01522		
Xi	126.8833	140.0785	235.3564	143.9995	142.8704	Average	=AVERAGE
						stdev	
						%	

Calculate the average and the standard deviation:

=average(P8:Y8)

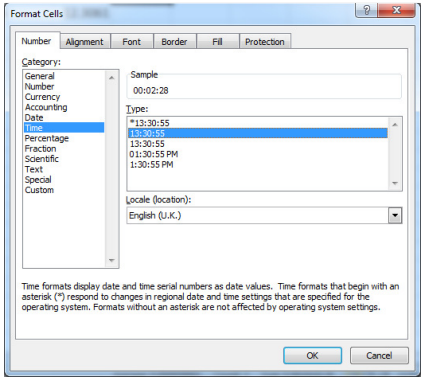
=stdev(P8:Y8)

And the Standard deviation as a percentage of the average (mean).

To work out the time in minutes and seconds simply divide the time in seconds by the number of seconds in 24 hours (86400):

Average	157.8376	=R8/86400
stdev	43.86977	
%	27.79425	

Now change the cell format to ‘Time’

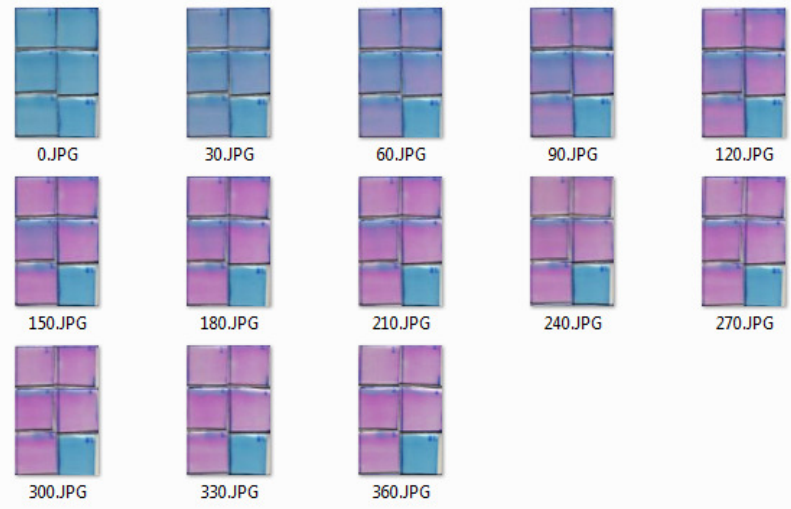


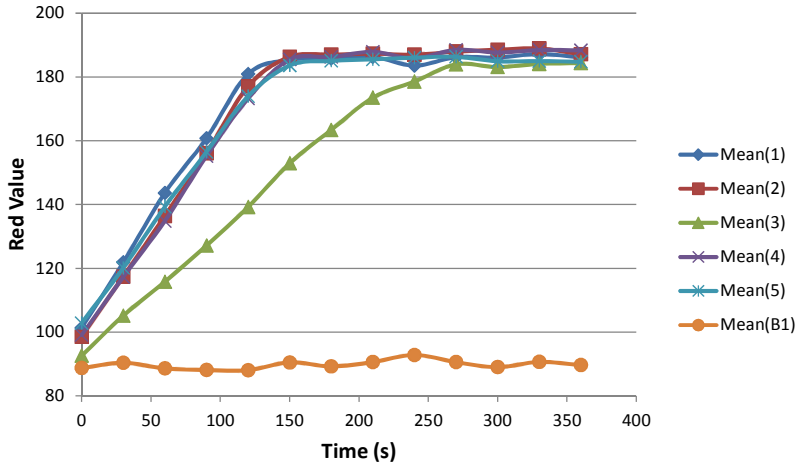
The result should look like so:

	Mean(1)	Mean(2)	Mean(3)	Mean(4)	Mean(5)			
c1	102.0474	97.9078	92.8784	98.6268	102.7906			
m1	0.659997	0.65237	0.384173	0.620783	0.595333			
c2	185.3962	190.2396	181.6768	187.8599	190.0211			
m2	0.003103	-0.00677	0.00688	0.001107	-0.01522			
Xi	126.8833	140.0785	235.3564	143.9995	142.8704	Average	157.8376	00:02:38
						stdev	43.86977	00:00:44
						%	27.79425	

Typical Results:

Pictures from a typical set of data:





	Mean(1)	Mean(2)	Mean(3)	Mean(4)	Mean(5)				
c1	102.0474	97.9078	92.8784	98.6268	102.7906				
m1	0.659997	0.65237	0.384173	0.620783	0.595333				
c2	185.3962	190.2396	181.6768	187.8599	190.0211				
m2	0.003103	-0.00677	0.00688	0.001107	-0.01522				
Xi	126.8833	140.0785	235.3564	143.9995	142.8704		Average	157.8376	00:02:38
							stdev	43.86977	00:00:44
							%	27.79425	