

BoxPlots in Excel 2007

Introduction

This page has been requested by Julie: here is what she said to me:

Hello Duncan,

Thank you so much for providing this information. Excel has so much packaged within it and it's not always easy to discover what you want. Question: can you update your box plot pages for EXCEL 2007, please? There is no longer a Wizard and I am having trouble following the instructions. The fact that EXCEL was not written for epidemiologists makes life just that bit harder.

With thanks,

Julie

From my initial page on BoxPlots, <http://www.duncanwil.co.uk/BoxPlot.html>:

A BoxPlot, or box and whisker diagram, provides a simple graphical summary of a set of data. It shows a measure of central location (the median), two measures of dispersion (the range and inter quartile range), the skewness (from the orientation of the median relative to the quartiles) and potential outliers (marked individually). BoxPlots are especially useful when comparing two or more sets of data.

<http://www.coventry.ac.uk/ec/~nhunt/BoxPlot.htm>

Despite the need for BoxPlots by many people, even Excel 2007 does not provide us with a template from which to plot them.

How to Draw a BoxPlot Chart

Suppose we have data from three groups, A, B and C for which we want to draw a BoxPlot.

Calculate the Statistics

Calculate the statistical functions **in this order**

=QUARTILE(,1)

=MIN

=MEDIAN

=MAX and

=QUARTILE(,3)

... as shown in the screenshot that follows:

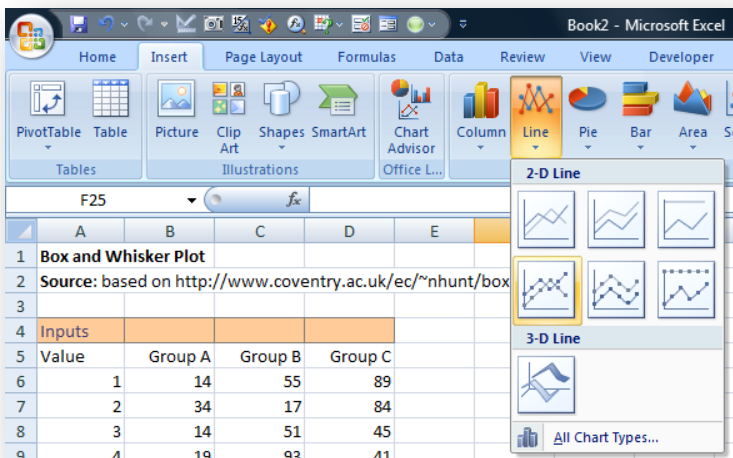
	A	B	C	D
1	Box and Whisker Plot			
2	Source: based on http://www.coventry.ac.uk/ec/~nhunt/boxplot.htm			
3				
4	Inputs			
5	Value	Group A	Group B	Group C
6	1	14	55	89
7	2	34	17	84
8	3	14	51	45
9	4	19	93	41
10	5	24	49	75
11	6	39	99	26
12	7	36	42	44
13	8	28	63	64
14	9	10	91	23
15	10	40	72	70
16	Statistic	Group A	Group B	Group C
17	q1	=QUARTILE(GrpA,1)	=QUARTILE(GrpB,1)	=QUARTILE(GrpC,1)
18	min	=MIN(GrpA)	=MIN(GrpB)	=MIN(GrpC)
19	median	=MEDIAN(GrpA)	=MEDIAN(GrpB)	=MEDIAN(GrpC)
20	max	=MAX(GrpA)	=MAX(GrpB)	=MAX(GrpC)
21	q3	=QUARTILE(GrpA,3)	=QUARTILE(GrpB,3)	=QUARTILE(GrpC,3)

Creating the Box and Whisker Chart

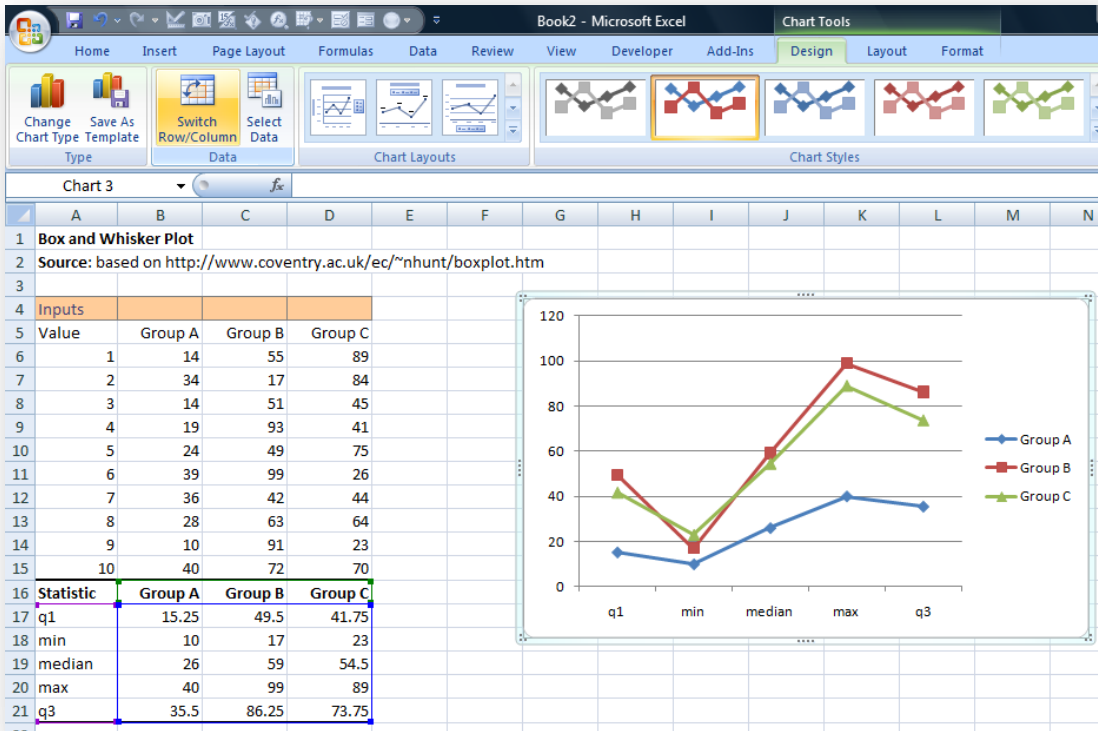
I am going to demonstrate two methods of preparing a box and whisker chart. One method gives relatively informal results and can be prepared as much for its looks as its information content. The other method gives results that are more traditional and can be used anywhere.

Method 1: using a Line Chart

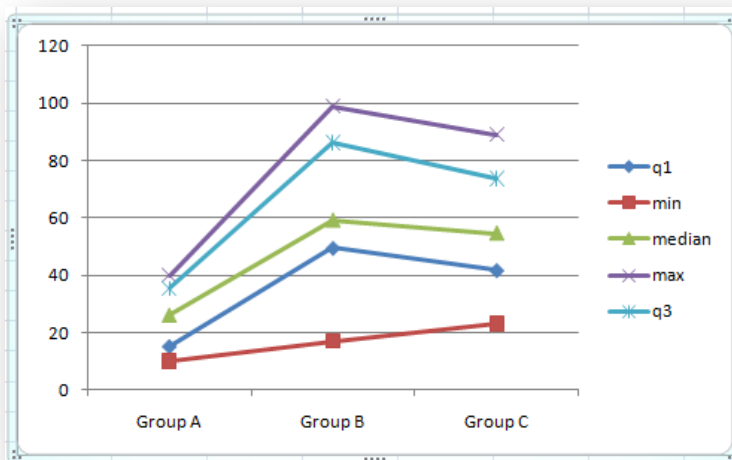
- 1 Highlight the area of the table containing the statistics: A16:D21, **including** the column headings and the values and ...
- 2 Select **Insert** from the main menu under Charts select a **Line** chart and choose the **Line with Markers** option.



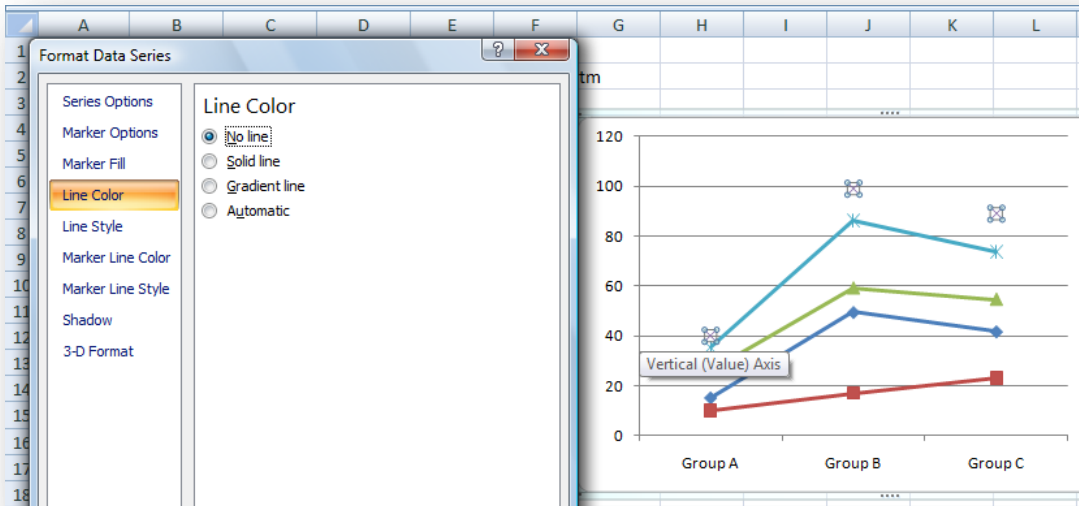
- 3 With your **chart selected** and under Chart Tools select **Design > Switch Row/Column**.



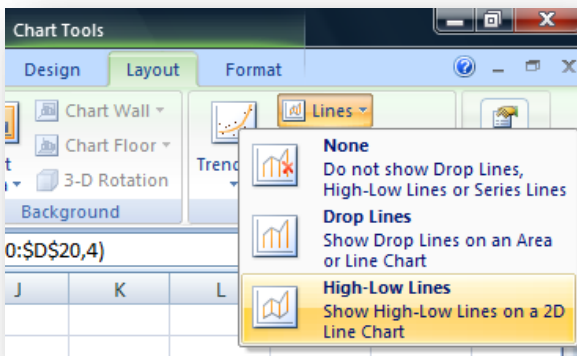
To give ...



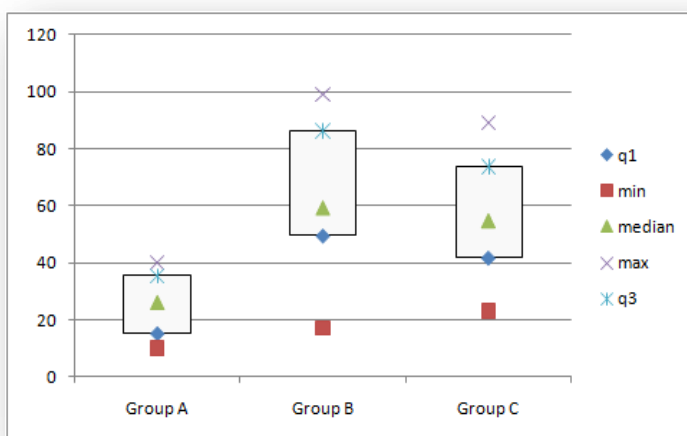
- Right-click on a data point from the first data series, and choose **Format Data Series > Line Colour > No line** to remove the connecting lines. **Repeat** for the other four data series in turn.



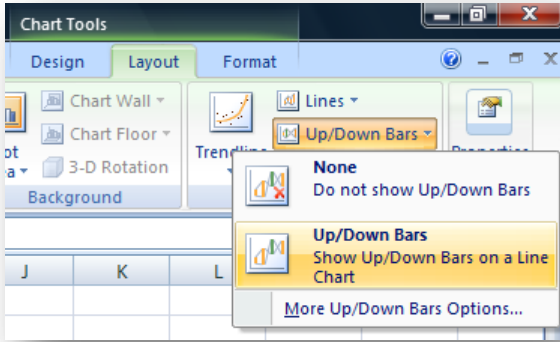
5 Select any of the data series and under Chart Tools select **Layout > Lines > High-Low Lines**



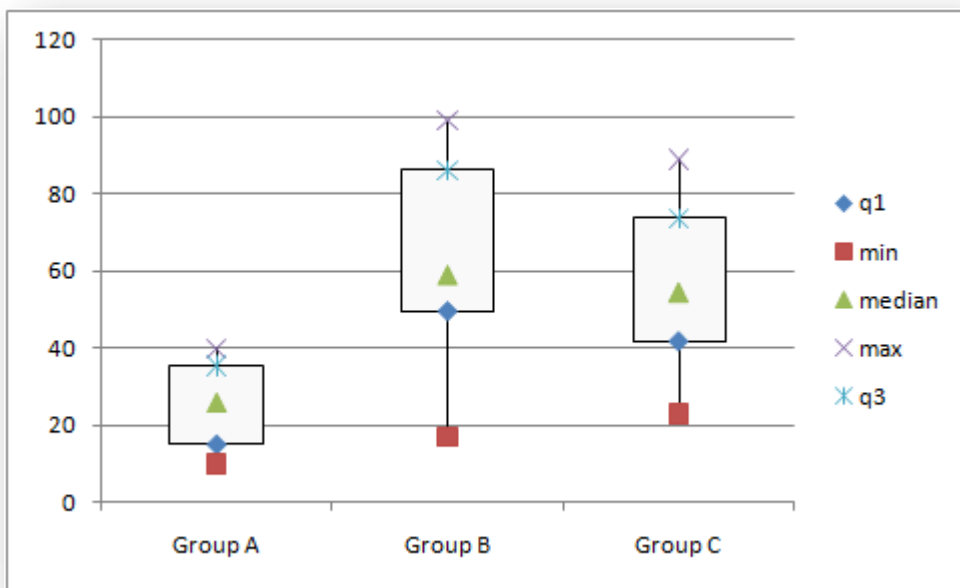
To give



then **Layout > Lines > Up/Down Bars > Up/Down Bars**

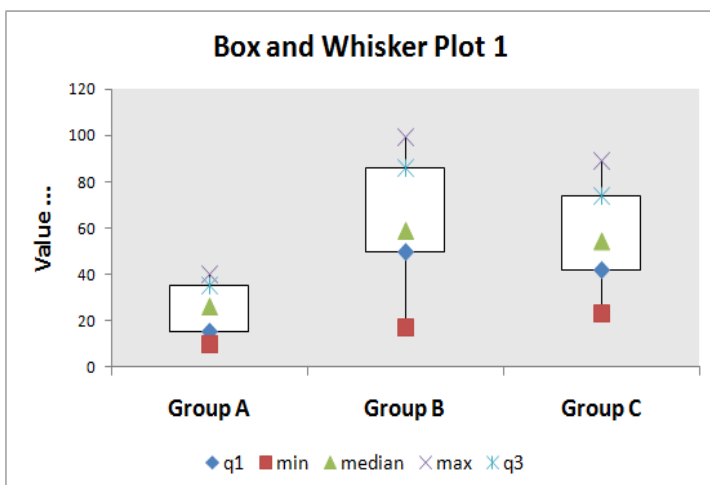


To give



- That's it! Further customising, such as adding titles and moving the legend, can be carried out according to your own preferences by right clicking on the relevant object and selecting the **Format** option on the shortcut menu.

Here is the finished BoxPlot for you:



Box and Whisker Charts

© Duncan Williamson July 2009 amended January 2010

Page 5 of 12

Method 2: using a Stacked 2 Column Chart

The first point to note here is that the order in which we prepare the statistics for this new BoxPlot is different from the previous method. Now it is:

max
q3
median
q1
min

The calculations and formulas are **initially** the same as before

	A	B	C	D	E	F
1	Box and Whisker Plot 2a					
2	Based on: http://www.bloggpro.com/box-plot-for-excel-2007/					
3						
4	Inputs					
5	Value	Group A	Group B	Group C		
6	1	50.5	46.7	62.5		
7	2	51.3	45.5	64.2		
8	3	55.3	45.6	66.1		
9	4	50.3	46.3	66.7		
10	5	55.0	49.7	63.4		
11	6	59.6	49.8	67.7		
12	7	51.3	48.5	62.2		
13	8	56.1	48.7	68.4		
14	9	59.7	48.8	62.7		
15	Statistic	Group A	Group B	Group C		
16	max	59.7	49.8	68.4		
17	q1	51.3	46.3	62.7		
18	median	55.0	48.5	64.2		
19	q3	56.1	48.8	66.7		
20	min	50.3	45.5	62.2		

but then:

Since we will be using a stacked 2D Column Chart we have to modify our data. These changes are only made to enable us to draw the chart properly.

- MAXIMUM value B16-B19 (that is the length of the top whisker)
- Q1 or 1st QUARTILE: copied from row 17 as it remains unchanged
- MEDIAN or 2nd QUARTILE: B18-B17
- Q3 or 3rd QUARTILE: B19-B18
- MIN: B17-B20 (that is the length of the bottom whisker)

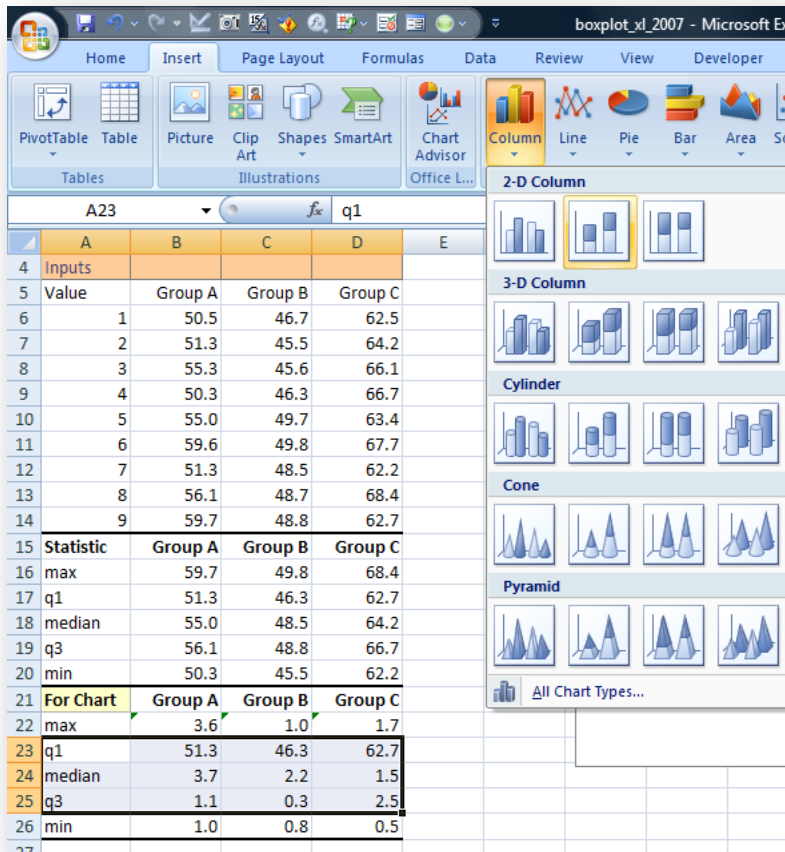
These values are calculated in the new range B22:D26

21	For Chart	Group A	Group B	Group C
22	max	3.6	1.0	1.7
23	q1	51.3	46.3	62.7
24	median	3.7	2.2	1.5
25	q3	1.1	0.3	2.5
26	min	1.0	0.8	0.5

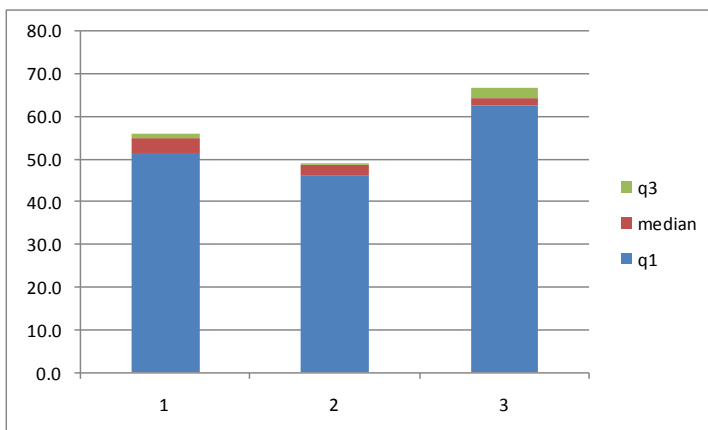
This method of preparing a BoxPlot uses the 2D Column Stacked Chart and we begin by:

Selecting the title, Q3, Median and Q1 results only, A18:D20, select

- Insert
- Charts
- Column
- 2D Column
- Stacked



This gives:



Acknowledgement: I am deeply grateful to Pallavi Mantha for pointing out that the first edition of this PDF file contained a chart that wasn't quite right! The above 2D stacked column chart is now the

Box and Whisker Charts

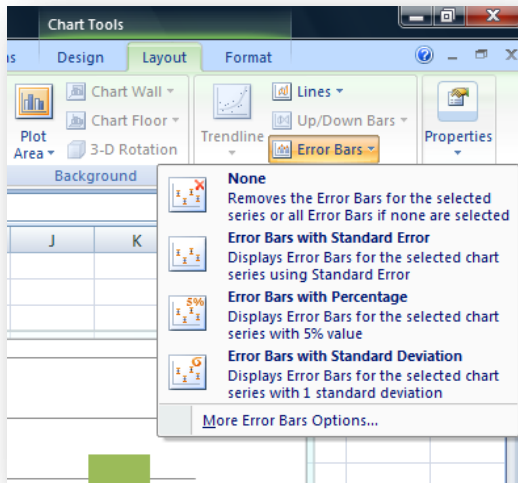
© Duncan Williamson July 2009 amended January 2010

Page 7 of 12

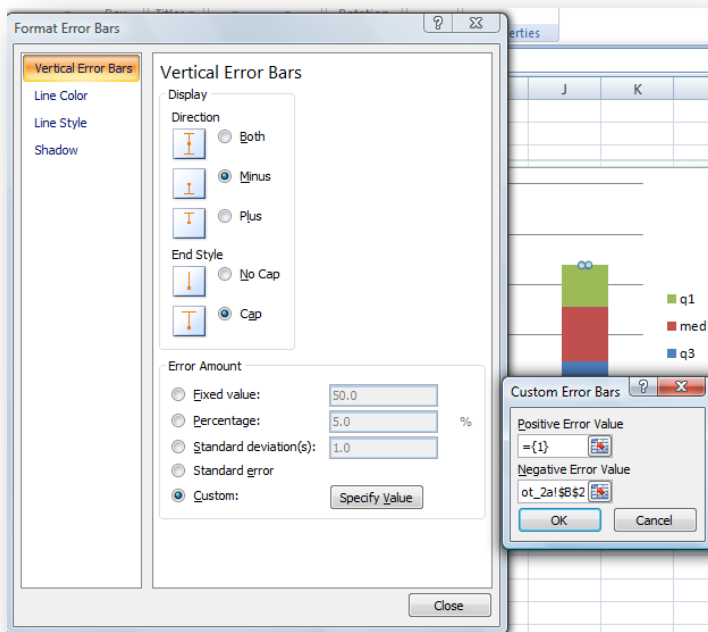
correct version. As always with these things, it is sometimes a mystery how they go wrong. Looking at the screenshot of what I did initially, I would swear blind that I have done the same again to correct the error! Obviously not, though!

Add the **minimum value whiskers** by doing the following, **with the Q1 bar selected on the chart:**

- Chart Tools
- Layout
- Select Error Bars

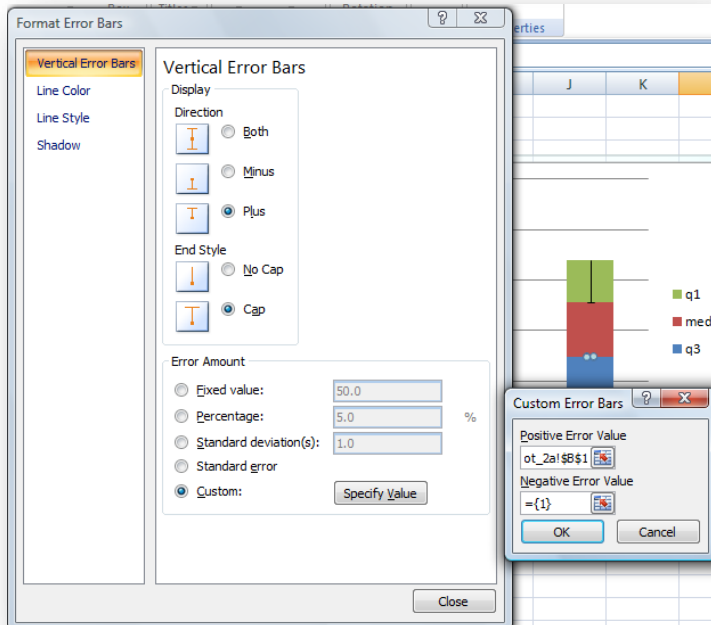


- More error bars options
- Select the Display Direction: Minus
- Indicate the Error Amount: Custom
- Click the Specify Value button
- Leave the Positive Error Value as it is
- Select the range B26:D26 for the Min values for the Negative Error bar



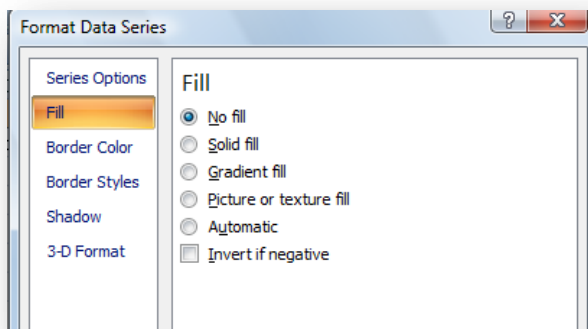
For the max value whiskers begin by selecting the Q3 bar on the chart

- More error bars options
- Select the Display Direction: Plus
- Indicate the Error Amount: Custom
- Click the Specify Value button
- Leave the Negative Error Value as it is
- Select the range B22:D22 for the Min values for the Positive Error bar



You need to make the Q1 bars invisible now:

- Click on any of the Q1 bars: any of them and they will all be highlighted
- Rick click
- Format Data Series
- Fill>No Fill
- Border Color>No Line

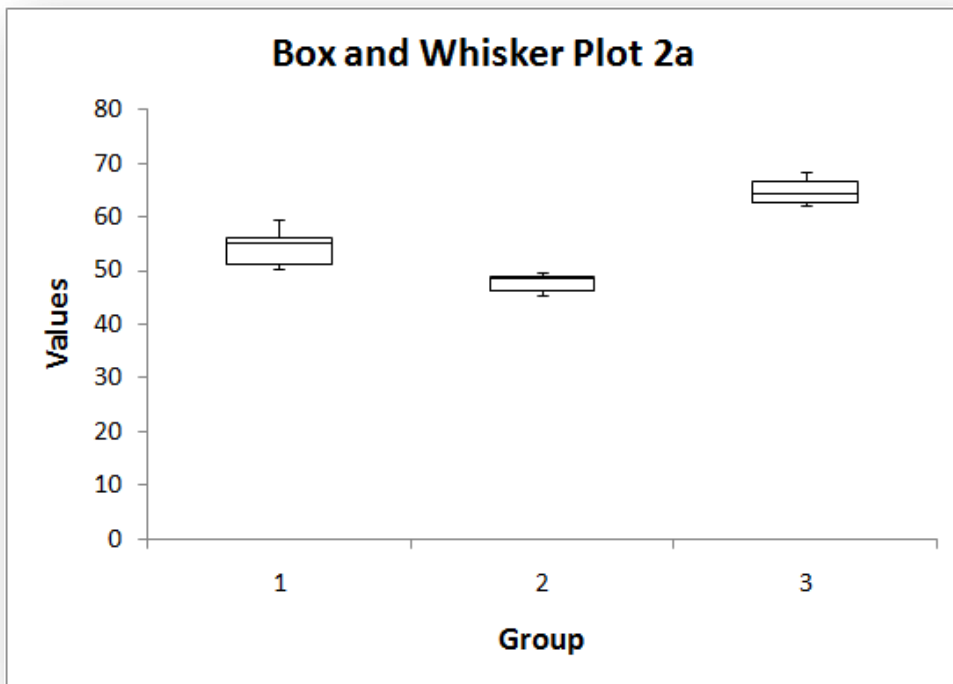


For the remaining two data series, format them to have

- No Fill
- Border Colour>Solid Line>Color>Black
- Delete the Legend

- Delete the Gridlines

The chart now should look like the one in the image below.



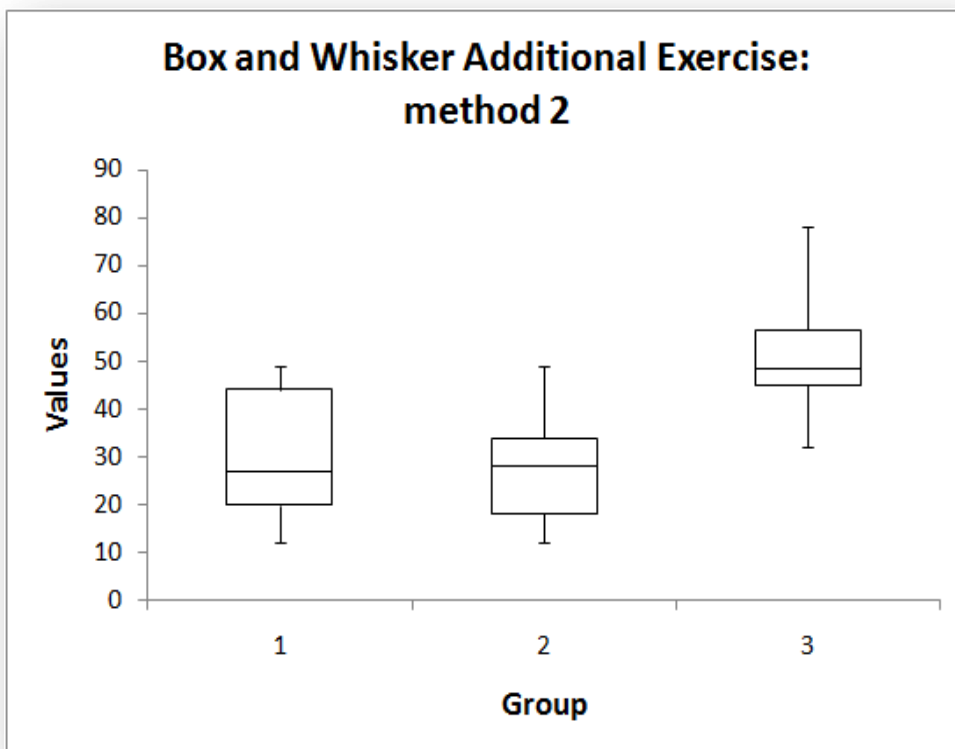
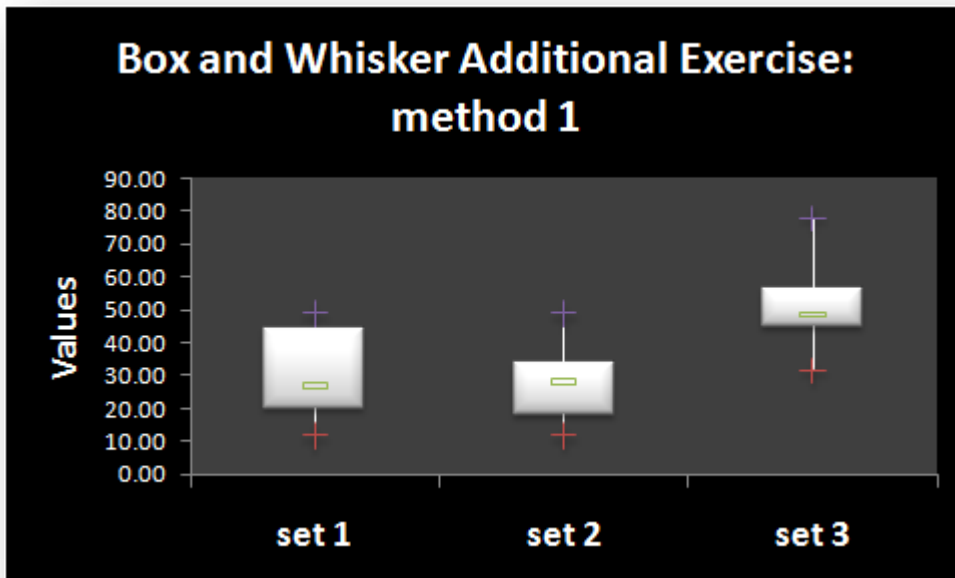
Additional Exercise: for you to try

Prepare a Box and Whisker Chart from the following data using methods 1 and 2 above. In addition to your basic chart, you should format the chart so that it has

- A meaningful title
- Meaningful axis labels
- Use Excel 2007's built in Chart Design features to enhance your chart as much as possible

Value	set 1	set 2	set 3
1	47	49	68
2	12	29	48
3	49	12	58
4	19	44	49
5	31	15	48
6	17	17	52
7	23	21	78
8	47	33	44
9	35	34	32
10	22	27	33

The following chart is my completed version using Method 1 but I have not shown any of my workings!



I have based **method one** of this page on Neville Hunt's page on this topic and the following are taken directly from there:

Acknowledgements from Neville Hunt's page

An earlier version of this article was published in The Spreadsheet User Volume 3, Number 2, November 1996. I am grateful to [Rodney Carr](#) for greatly simplifying my original method. If you find this article useful, why not subscribe to [Teaching Statistics](#) magazine, where I regularly contribute articles on drawing statistical charts using Excel.

References

Daly, F, Hand, D J, Jones, M C, Lunn A D and McConway, K J (1995). *Elements of Statistics*. Addison Wesley /The Open University.

Devore, J and Peck, R (1990). *Introductory Statistics*. West Publishing Co.

The address of Neville's page is: <http://www.coventry.ac.uk/ec/~nhunt/BoxPlot.htm>

I have based **method two** of this page on this page <http://www.bloggpro.com/box-plot-for-excel-2007/> but I have to say that I had a lot of work to do to get this method to work properly as the instructions on that page did not work!

Other References used in this chapter:

Les Oakshott (1994) *Essential Elements of Business Statistics* DPP