

How do you do that?

A small case using a couple of useful Excel functions

The functions:

The functions that we shall find particularly useful in this case are the =COUNTIF() function, and the =INDEX() function. In addition, to tidy things up at the end, we will use a tool from the Forms toolbar.

COUNTIF

=COUNTIF(*range,criteria*) returns the count of non-blank cells in the *range* that meet the *criteria*

INDEX

=INDEX(*table, rownum, colnum*) looks in *table* and returns the value at position *rownum, colnum* in *table*.

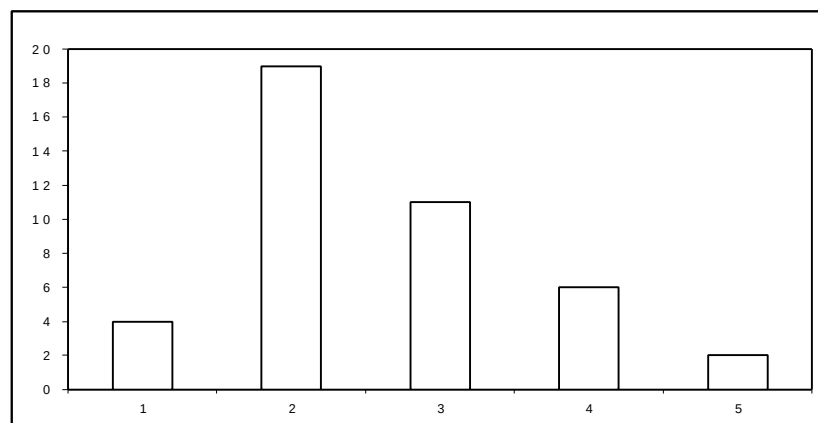
The Situation

The data in this case arise from an internal “Quality Culture” survey at a manufacturing and distribution company. The survey was distributed to all employees, and 43 of the 50 employees returned the survey. There were 73 quantitative questions in the survey, of the Likert type, a scale of 1-5, with 1 reflecting poor performance, and 5 indicating superb performance, while 3 was a neutral score. There were also a number of qualitative variables, with which we will not concern ourselves for now.

There are many sensible numerical and graphical summaries that could be used to help convert these data to information. We will focus on the one graphical tool, the bar-chart, and will also use a single numerical performance measure. With 73 questions to analyse (before we even begin to think about the possible inter-relationships among these variables), there is potentially a great deal of graphing to be done! Clearly, this is a case where some time spent in designing the tools to be used in the worksheet will pay off when the final analysis is done.

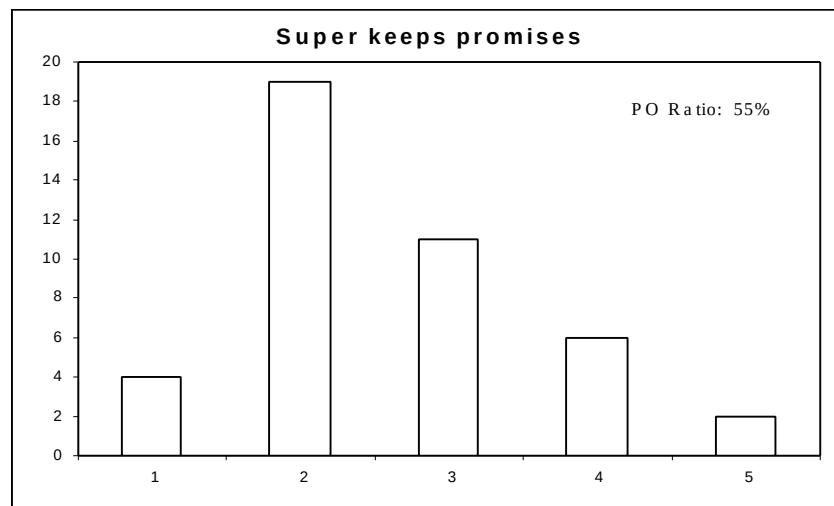
The design.

First, the chart. We have stated that a bar chart will be the most appropriate. There will be five categories to be listed along the bottom of the bar chart, the numbers 1 to 5. For each category, we will plot the number of responses that fell into that category. In keeping with the usual guidelines for clear graphical display, we will create this bar chart in black and white, with a “no-frills” approach. An example is shown below.



In addition, we would like each chart to have a sensible title, that tells us (reasonably minimally) which variable is being graphed. Further, we should like some numerical summary to tell us whether this is a chart that “demands action” or is “just another chart”. For this we could choose to place the average response on the chart, with a standard deviation, or perhaps the median. While these choices are straightforward and commonplace, a stronger signal may be received if we take a “weakness focus”, and regard all those responses of 3 or less as being in some sense “negative”. This suggests that we might take

the percentage of responses of 3 or less as being a “dissatisfaction” score for any given question. Thus we might hope to get a chart something like this:



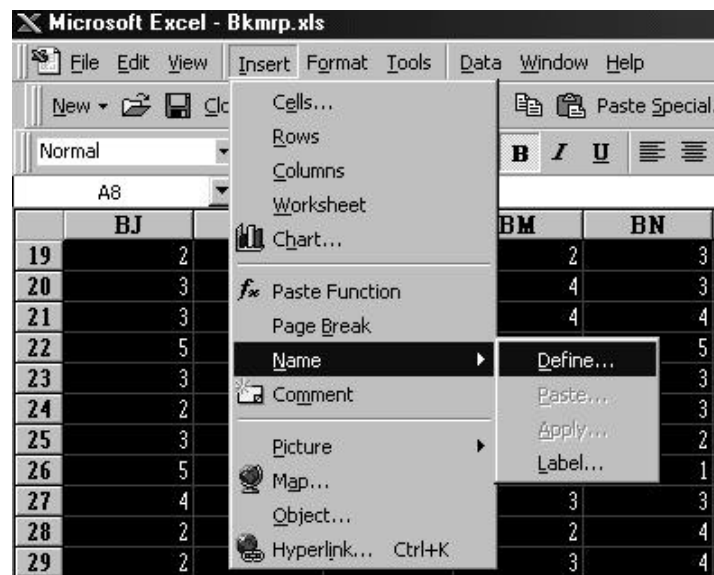
The “PO Ratio” of 55% indicates that 55% of respondents scored 3 or less on this question. The title “Super keeps promises” is a shortened version of the full question, which was “I have confidence my immediate supervisor generally keeps the promises he/she makes”.

The Spreadsheet

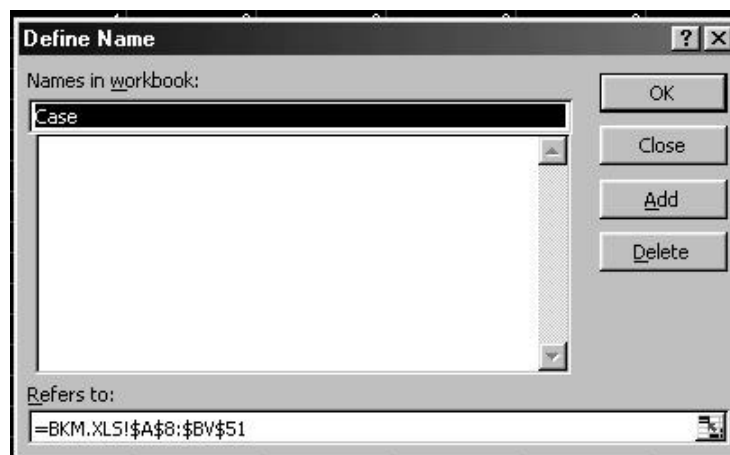
The full dataset contains 43 cases, and 73 variables, and lies in the range from A8 to BV51 (including the variable names Q1 to Q73, and the additional column for the case numbers). A small section is shown below.

8	Case	Q1	Q2	Q3	Q4	Q5	Q6
9	1	2	2	2	4	4	4
10	2	3	3	1	3	2	2
11	3	2	2	4	3	2	2
12	4	4	4	2	3	2	4
13	5	3	4	4	5	3	4
14	6	4	2	2	3	4	3
15	7	4	5	3	3	5	3
16	8	2	2	2	2	2	1
17	9	3	4	4	3	2	2
18	10	2	2	2	2	2	2
19	11	4	1	2	3	3	1
20	12	2	3	2	2	1	3
21	13	4	3	3	3	4	3
22	14	2	4	3	3	4	4
23	15	4	3	3	2	3	3
24	16	4	2	3	3	5	4

The first step we will take is to name this entire area as **Data**. We can highlight the block of data by placing the cursor on cell A8 (as in the figure above) and then, holding down both the control key and the shift key, pressing the right arrow key followed by the down arrow key. To name the area, we select the **Name** option from the Insert menu, and then the **Define** option, as shown in the figure below:



This results in a dialog box as shown:



The highlighted bar contains the word “Case” – this is simply because this is the text entry at the top left of the area in the worksheet. Type **Data** in here, and click on the OK button to complete defining the area.

Creating a chart

Since we want to create a chart in a “reusable” fashion, we will create a “work area” at the bottom right of the area called **Data**, beginning in cell BW52. (We can get there quickly using the **Go To** command, found on the **Edit** menu, or use the Ctrl-G combination, or by using the control and arrow key combination to move quickly to the bottom right of the data block, and then moving to cell BW52.) In cell BW52, enter **Work**, and then name this cell using the procedure above. This time, we can simply use the default name – the contents of the cell. Beneath this cell, (i.e. beginning in cell BW53) we will place a sequence of numbers, from 1 to 44 (there are 44 rows in the data area, including the titles row). We can do this by starting the sequence and then using the fill handle to extend it, or by placing a formula in these cells and extending this down as far as is required.

We will begin by copying the first column of the data area into the region at **Work**. Take the data from the title Q1 down to the last entry in that column, copy it (Ctrl-C) and place it alongside the index numbers 1-44 just entered. A small portion of the spreadsheet, showing some of the results of the work just carried out is shown below.

2		
3		
	Work	
	1	Q1
	2	2
	3	3
	4	2
	5	4
	6	3
	7	4
	8	4
	9	2

At this stage, the data we are working on is placed in the area BX53 to BX96. We will create a chart for these data, along the lines of our design ideas. When this is complete, we can easily create a chart for another column of the data by copying the data and pasting it into this area. (We will do better than that eventually, but this is the first step along the way to creating a more flexible tool.)

Below the column of data, we will be working on creating the chart – we will begin this part by naming a cell as **ChartWork**. Below the working data, in cell BW98, type **ChartWork**, and then name the cell using the procedure already described. In the cells immediately below the named cell, place the sequence 1 – 5, descending down the worksheet (cells BW99:BW103). These are the possible values that the variable can take – we need then to count the number of occurrences of each in the working data area above. For this we use the =COUNTIF() function. The format of the =COUNTIF() function is as described in the opening paragraphs, and repeated below:

=COUNTIF(*area,criterion*) scans through the *area* and returns the number of times that the *criterion* was met.

We want to count the number of occurrences of the digit '1' in the area corresponding to the column of data that we are currently working on, and then the number of '2's and so on. The data are in the area BX53 to BX96, so the appropriate formula is as shown in the display below.

95			43	4	
96			44	1	
97					
98			ChartWork		
99			1	=countif(\$BX\$54:\$BX\$96,BW99)	
100			2		
101			3		
102			4		
103			5		
104					

Note that the area to be scanned has been made into an absolute address range – this is so we can simply drag the formula down, either by using the fill handle, or by highlighting the area to be filled and then using the Fill Down (Ctrl-D) command from the Edit menu. The criterion is a simple one in this case, it simply tests to see whether an element is equal to the contents of the cell at left. (In the formula above, this is cell

BW99. This is a relative address, since we want it to change as we drag the formula down the column.) The result of this is shown below:

	43	4
	44	1
	ChartWork	
1		2
2		13
3		14
4		14
5		0

This, then, is the set of data that we want to drive the bar chart. We might hope that simply highlighting the area from just below **ChartWork** to the right and down to the cell containing the zero and then using the **Chart Wizard** would generate the appropriate chart. Alas, this is not the case! Excel expects the labels on the horizontal axis (The Category axis) to be strings, not numbers. We need to resort to some artifice to accomplish this, by turning these numbers (the sequence 1 – 5) into strings. We can accomplish this by concatenating a space character and the contents of the appropriate cell. (the ampersand character, &, is used for concatenation). This is shown in the display below, where the formula for the labels is shown.

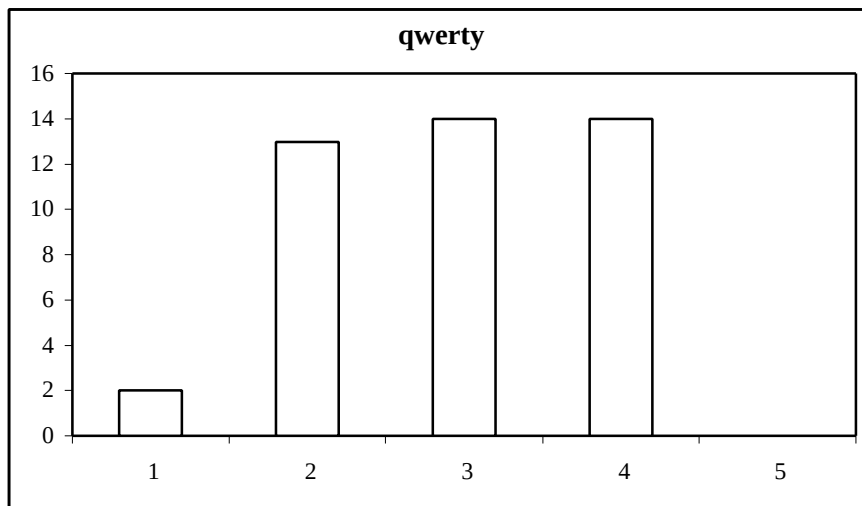
ChartWork		Labels
1	2	= " "&BW99
2	13	
3	14	
4	14	
5	0	

If we place another column to the right of this Labels column, and simply map the counts across, we get the following:

	BV	BW	BX	BY	BZ
91		39	2		
92		40	2		
93		41	3		
94		42	4		
95		43	4		
96		44	1		
97					
98		ChartWork		Labels	Counts
99		1	2	1	2
100		2	13	2	13
101		3	14	3	14
102		4	14	4	14
103		5	0	5	0

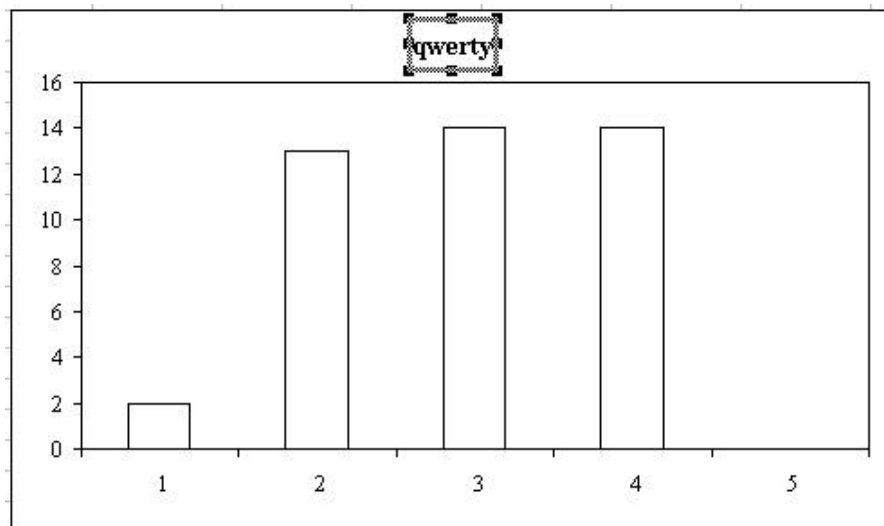
The data in the area BY99:BZ103, then, can be put into the **Chart Wizard** to create a bar chart (Note that Excel calls this a “Column Chart”). The chart, of course, needs to be “cleaned up” a little to make a reasonable presentation. The chart below has had the Legend removed, the background colour removed,

and the colour within the bars has been removed. The title of the chart has been entered as a “nonsense” string. We will replace this with something meaningful, driven from within the spreadsheet.



Putting on the title

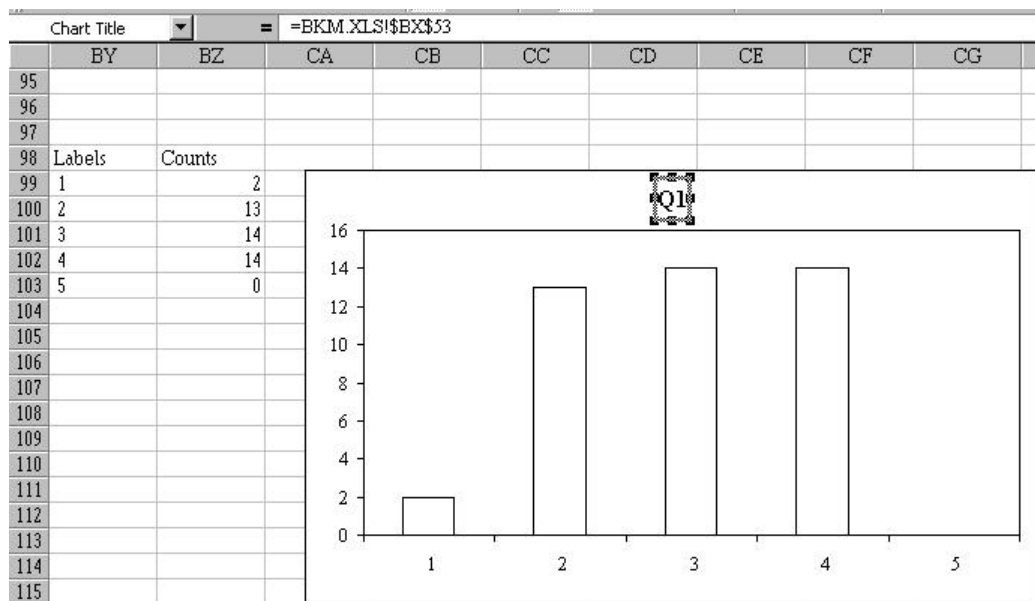
The title in the chart above was entered as part of the **Chart Wizard**. We could have entered any string, but we'd like the title to be relevant to the data that have been copied into the work area. Since we copied the title at the head of the column as well, we should be able to use this to “drive” the title of our chart. To accomplish this, simply click on the title of the chart. A box will appear around the title, with resizing handles around it. We can now click with the mouse in the editing bar



And type the formula

`=BKM.XLS!$BX$53`

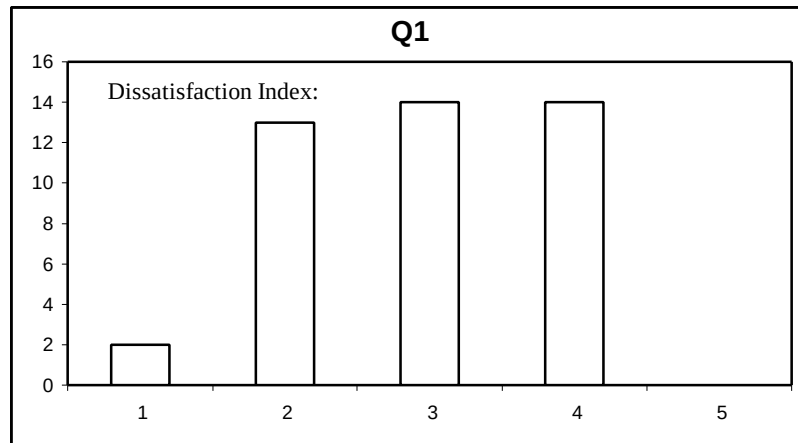
(It is easier, of course, to simply type the equals sign and then “point” at the appropriate cell with the mouse pointer.) The title of the column of data is in cell BX53, at the top of the set of data that has been copied down.



What further embellishments do we want for the chart? As part of our design, we decided that we would like the proportion who stated that the company's performance was at best neutral should also be displayed, and we called this a "dissatisfaction index". The summary statistics we need to calculate this index are easily obtained. We need the total of the counts corresponding to values 1,2 and 3, and of course, the grand total of respondents for the question. The =SUM() function can easily give these two sums, and then our "dissatisfaction index is simply their ratio. We can then express this ratio as a percentage (with no decimal places), and concatenate the result with the string "Dissatisfaction Index: ". This is shown below.

	BU	BV	BW	BX	BY	BZ
95			43	4		
96			44	1		
97						
98			ChartWork		Labels	Counts
99			1	2	1	2
100			2	13	2	13
101			3	14	3	14
102			4	14	4	14
103			5	0	5	0
104						
105			Sum <=3	29		
106			Overall sum	43		
107						
108			Index			
109			(rounded)	67		
110						
111			="Dissatisfaction Index: "&BX109&"%"			

We can annotate the chart with this string, to give a statistical summary of the data that are currently under consideration. Click on the box containing the chart, so that the outside frame has its handles highlighted. Now type in the formula bar an equals sign, then point to the cell containing the "Dissatisfaction Index". Press the Enter key to obtain the chart below.



The title and annotating statistics make the display a little more useful for us, but remember, we have 73 variables, and the possibility of us remembering what each of these variables refers to is rather remote! We therefore need a more informative title. We can't use the whole of the question as a title, as this would give us a chart that was mainly title. However, there is a set of "abbreviated titles" in the worksheet, beneath the main body of the data (in the range I56:I128). Below is the first portion of the list, with the corresponding questions from the survey.

We have been given a clear picture of where BKM is going	Clear picture
BKM is succeeding at making our paperwork procedures simpler	Simpler paperwork
BKM is making the changes necessary to compete effectively	Necessary changes
BKM's emphasis on quality has resulted in measurable improvement in our services	Quality making improvements (Service)
I feel valued as an employee of BKM	Feel valued
Management's decisions about employees are usually fair	Fair decisions
BKM places more importance on quality of products or services than on cutting costs.	Quality before cost cuts
Individuals are recognised for their contributions to quality.	Individuals recognised for Q
BKM places more importance on quality of products or services than on meeting deadlines	Quality before deadlines
BKM's emphasis on quality has resulted in measurable improvement in our products	Quality making improvements (Product)
I have confidence in the decisions made by my upper manager	Confidence in upper manager
My upper manager communicates effectively	Upper mngr communicates

We shall name the area containing the abbreviated names (i.e. from I56 to I128) as **ShortNames**. (Highlight the area, then use Insert ... Name ... Define, and enter the name **ShortNames** in the dialog box.) The first name in the list is "Clear picture"; the second is "Simpler paperwork", .. the tenth is "Quality making improvements (Product)" and so on. Obviously, if we can look up the 1st, the 2nd, 10th item on a

list, surely Excel can be persuaded to do the same thing! And, indeed, this is the case. This is where the =INDEX() function comes in. The format of the function is (as stated at the beginning of this case):

INDEX

=INDEX(*table*, *rownum*, *colnum*) looks in *table* and returns the value at position *rownum*, *colnum* in *table*.

In this case, the *table* is the named area **ShortNames**. This consists of 73 rows, and one column. Thus the entry in row 1, column 1 is “Clear picture”. We can refer to this using the Excel function statement:

=INDEX(ShortNames,1,1)

Furthermore, we can always pick out the entry in **ShortNames** corresponding to any number N by the statement

=INDEX(ShortNames,N,1).

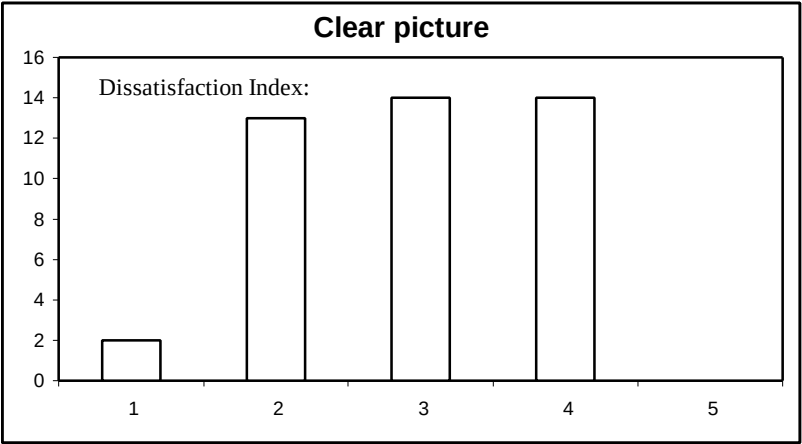
In our case here, we would like to use this function to pick out the appropriate short name for the current graph. Which graph is that? Well, the variable number is on the end of the variable name (Q1 etc), which is in cell BX52. For now, we will put a ‘1’ in cell BY52, and then, in cell BY53, we will put the index statement which will “pull out” the appropriate name from **ShortNames**. In cell BY53, type the formula:

=INDEX(ShortNames,BY52,1)

Below is this area of the worksheet, just after the formula has been entered.

2				
3				
Work				1
	1	Q1		Clear picture
	2		2	
	3		3	

The title of the column of data is currently in cell BX53, at the top of the set of data that has been copied down. We can simply click on the title in the chart, and change this address to BY53, and the title will change to the short name version. This is shown in the chart below:



Let Excel do the Copying

Just as Excel can copy the name across for us, so too can Excel copy all of the data down into the work area – a great deal faster and more effectively than we can using the mouse. Since the =INDEX() function was so useful in getting the name across, we might think it could come in handy here as well – and it does.

Remember that we named the original block of data as **Data**. As with the short names, we can extract any specific cell within this block of data by using the =INDEX() function. Notice that the 1st variable, Q1, is in the 2nd column of the block of data, and so on up to the 73rd variable, Q73, which is in column 74 of **Data**. Thus, if we want to extract the 1st variable using =INDEX(), we will have to pick out the second column:

=INDEX(Data,1,2)

=INDEX(Data,2,2)

=INDEX(Data,3,2)

=INDEX(Data,4,2)

and so on.

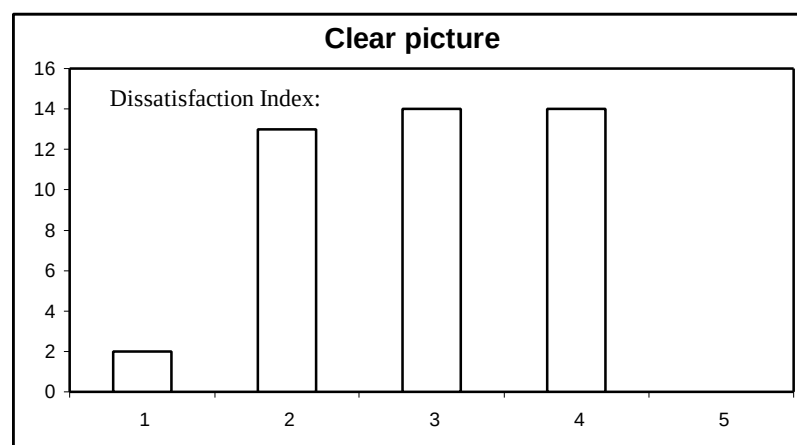
We can do this in the position where we have our working copy of the column of data: in cell BX53 type the formula

=index(Data,BW53,\$BY\$52+1)

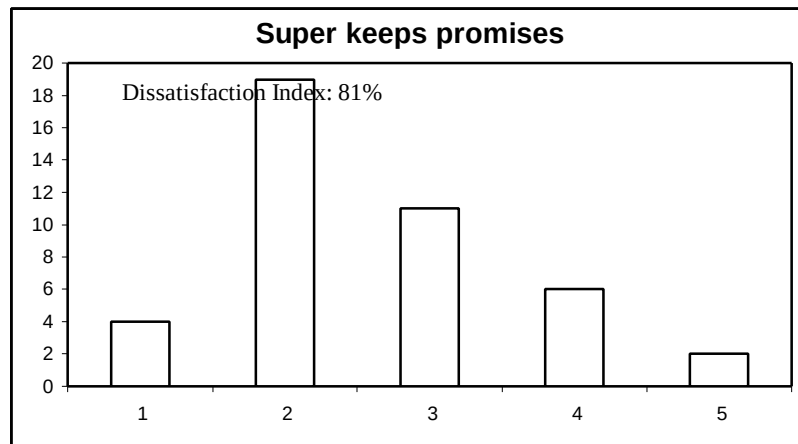
Note that the cell BW53 is a relative address – as we drag this formula down the column, we want to change the row that is being looked up each time. The cell BY52 has been made into an absolute address (press function key F4) because we want always to be picking out the variable that is defined by the contents of BY52.

	BV	BW	BX	BY	BZ	CA
49		4				
50		2				
51		3				
52		Work		1		
53			1	=index(Data,BW53,\$BY\$52+1)		
54			2	2		
55			3	3		
56			4	2		
57			5	4		

If we press the Enter key, and then copy this formula down the column (can use a double-click on the fill-handle to do this, or Ctrl-Shift-Down arrow and then Ctrl-D to fill down), we get the column of data as before. The chart at this point looks like this:

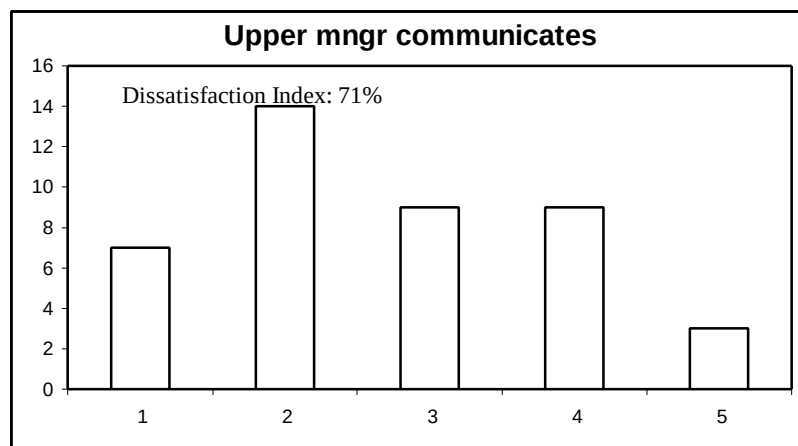


At this point, the charting tool is quite versatile and reasonably general: Everything is driven by the one cell, cell BY52. If we change the number in BY52, we choose another variable. Try it! Below is the result of placing 17 in cell BY52:



Note that we may still have a little work to do, as the statistical summary may have been placed in an inconvenient location. We can move this wherever we like using the mouse pointer.

When BY52 is changed to 12, we get the graph below:

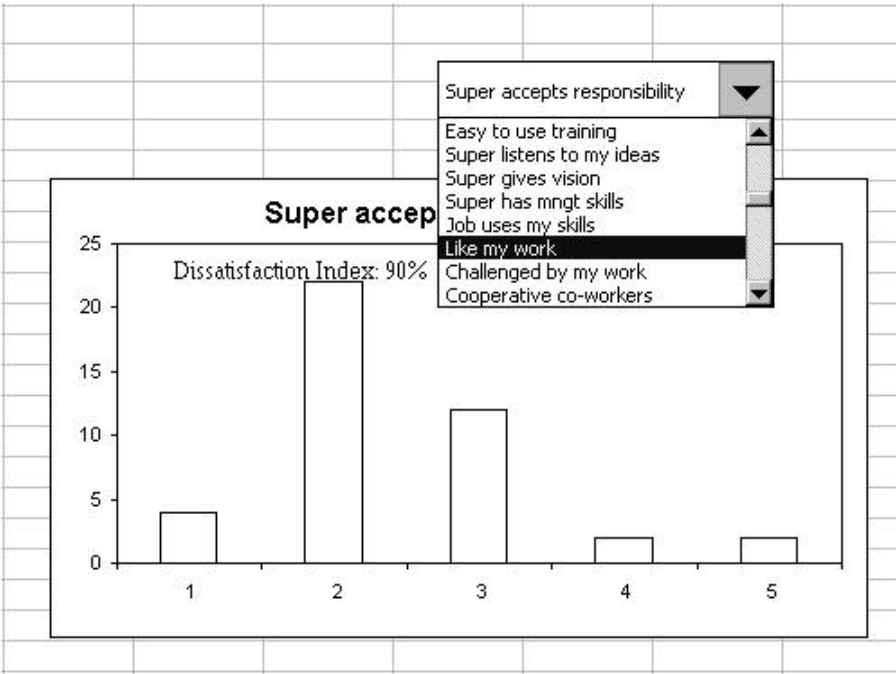


Thus our time spent in design and development can now pay off – we can browse through the 73 variables quite quickly, although this still involves a little paging around the worksheet. The final section gives us a slightly more elegant way of doing this.

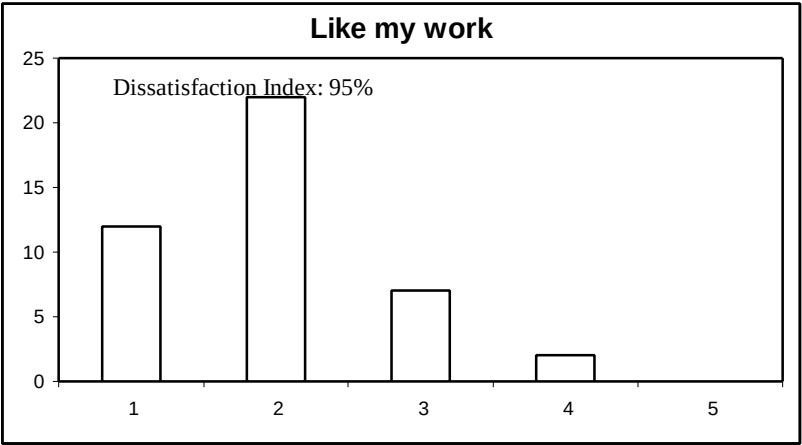
Putting on “The Ritz”!

We are used to using pull-down menus as a way of selecting actions when using windowing applications on personal computers. In our present case, this is the sort of action that we are carrying out – selecting a variable from a list (in this case) of 73 possible variables. Surely, again, we can get Excel to do the work? Again, the answer is Yes! The tool we shall use for this purpose is the “**Combo Box**”, a tool which is available on the **Forms** toolbar. When we implement the tool, we will be able to select from a list exactly which of the 73 variables we wish to view, and select by name, not by Q1, Q2 etc. Below is a view of

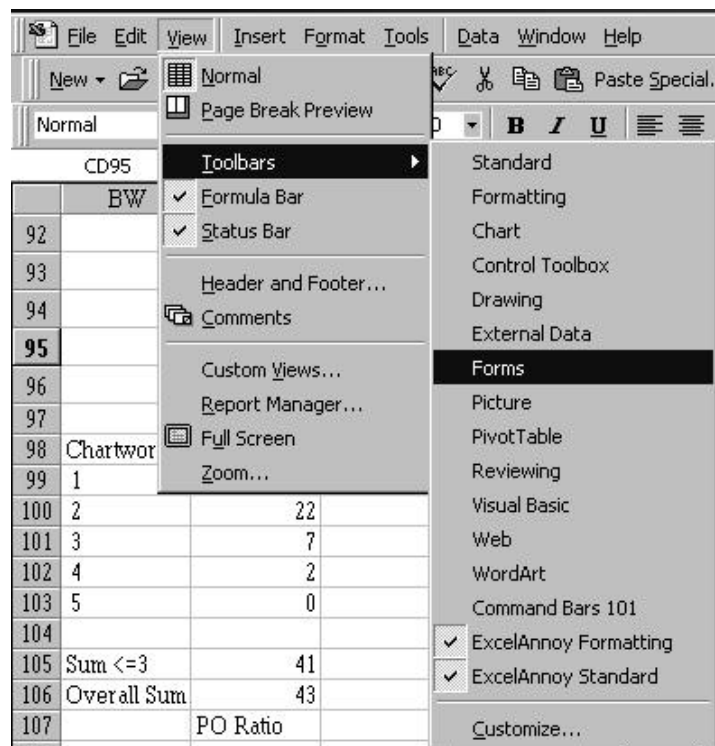
what the tool will look like in action: The current chart corresponds to “Super accepts responsibility”, and the chart about to be selected is “Like my work”.



The resulting chart is shown below:



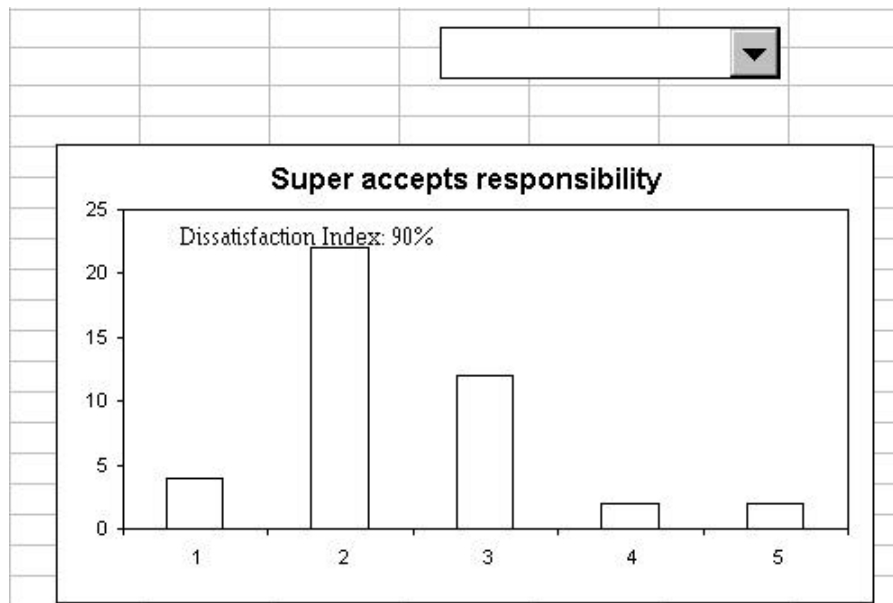
To put this tool in place is reasonably straightforward. First, we need to have the **Forms** toolbar available. Click on **View ... Toolbars ... Forms**, as in the display below:



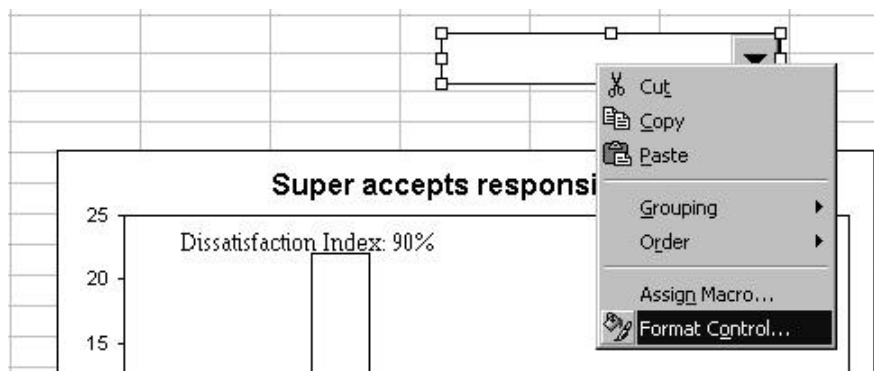
The result of this is the display of the **Forms** toolbar, shown below. The icon on this toolbar which appears slightly depressed is the **Combo Box**, as shown. This is the tool we will use. Click on the **Combo Box** icon, and then move the mouse pointer to cell CD95 (approximately – the combo box isn't actually located precisely in a range of cells, but rather sits "over" them, wherever you put it). Notice that the mouse pointer changes to quite a fine "cross-hair". Click and drag, as though you were drawing a rectangle.



The result of this should look something like this:



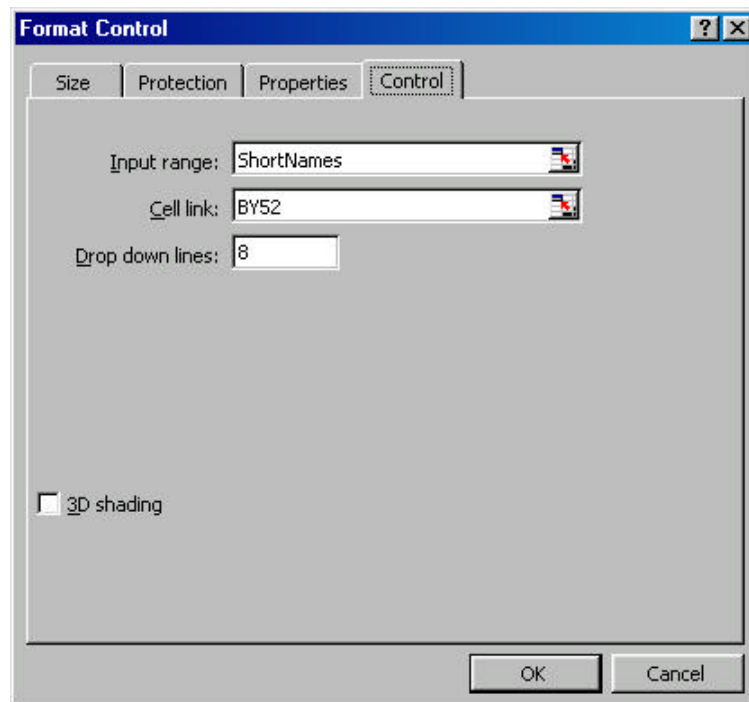
The blank rectangle with the arrow at the right-hand end is the Combo Box. At this stage it has no parameters to drive it. We want to “connect” it to the list of names to choose from, and to the cell that drives all our calculations and the chart. We do this through the “Format Control” of the Combo Box. To do this, right-click on the Combo Box, to get something like this:



The “Format Control” option is highlighted. If we select this option, we get this display:

The image shows the 'Format Control' dialog box in Microsoft Excel. It has four tabs: 'Size', 'Protection', 'Properties', and 'Control'. The 'Control' tab is selected. Inside the 'Control' tab, there are three input fields: 'Input range:' with an empty text box and a selection icon; 'Cell link:' with an empty text box and a selection icon; and 'Drop down lines:' with a text box containing the number '8'. Below these fields is a checkbox labeled '3D shading' which is currently unchecked. At the bottom right of the dialog are 'OK' and 'Cancel' buttons.

The input range is the range that will be used as the list to display in the **Combo Box** and choose from. The **Cell link** tells the **Combo Box** where to put the “result” of the choice. This result is simply the order in the list of the choice. Thus the first item on the list returns the value 1; the 17th item returns the value 17 and so on. “**Drop down lines**” dictates how long a list is displayed at a time. For our purposes, of course, the list is in the named range **ShortNames**, and we want the result to go into cell BY52, as this is the cell that changes the work area and the chart. Here is the **Format Control** dialog, filled out.

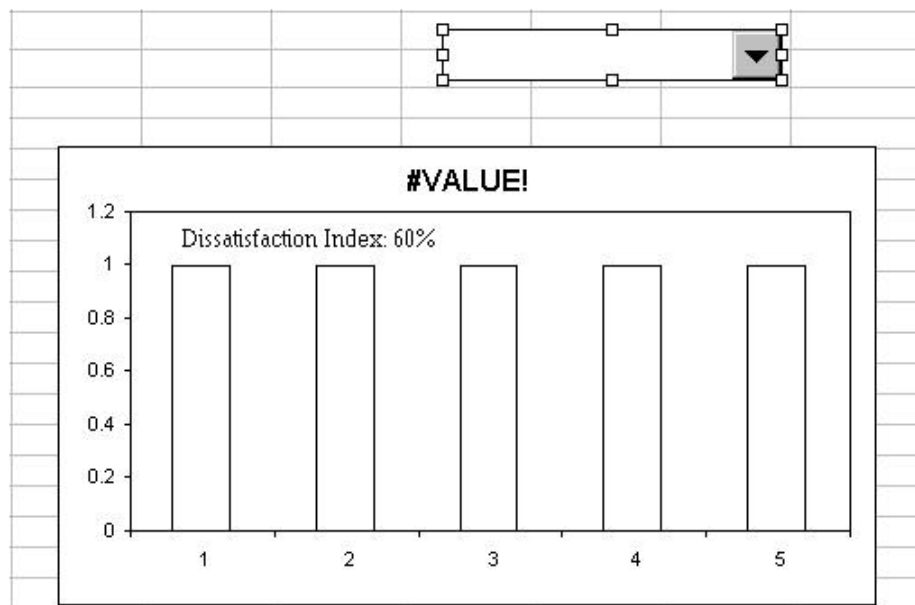


The **Format Control** dialog box is shown with the **Control** tab selected. It contains the following fields:

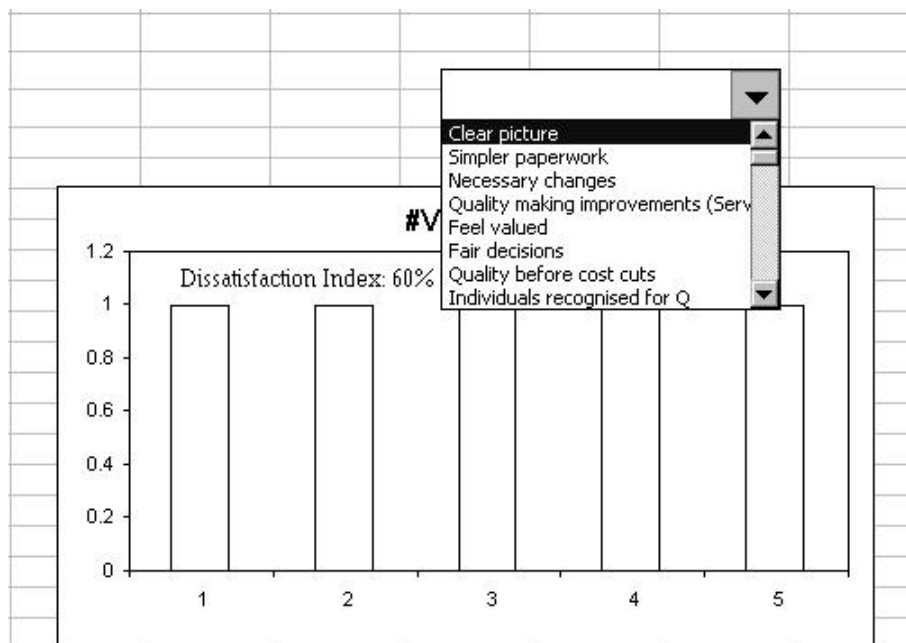
- Input range:** ShortNames
- Cell link:** BY52
- Drop down lines:** 8
- ☐ **3D shading**

At the bottom are **OK** and **Cancel** buttons.

If we click the OK button, this establishes the format control for the **Combo Box**, but does not initialise any value for the box. As a result, clicking OK will give something like this:



Click in a cell of the spreadsheet (to deselect the Combo Box), and then click on the **Combo Box** arrow, and select the first variable listed:



The result is:



The tool is now a very versatile tool for browsing through the complete list of 73 variables. Try it!

A final word. When copying pictures from here into a Word document, it pays to copy as a picture (Hold down the shift key as you click on the **Edit** menu, and select "**Copy Picture**"), and then to paste into Word using **Paste Special ... as Picture ...** and uncheck the "**Float over text**" option. This will result in smaller files that are easier to control.