Basic Excel Competencies

Below is a list of basic spreadsheet tasks which you should feel comfortable with.

- Select a range of cells, by “click and drag” or by using the Shift and “Arrow” keys.
- Enter text or numerical data into a cell.
- Move the contents of a cell into another location.
- Correct incorrect information in a cell.
- Erase the contents of a cell.
- Erase a whole set of cells at once.
- Use the **Undo** command as needed.
- Discard a worksheet.
- Open a new worksheet.
- Save a worksheet.
- Retrieve a worksheet.
- Use the fill handle to create a sequence of months, days, dates, times, or numbers.
- Centre text across more than one column.
- Use the alignment icons to align text and/or numeric information in a cell.
- Build a formula to add, subtract, multiply or divide into a cell. Use absolute and relative addresses.
- Use the Autosum icon to sum a contiguous range of cells.
- Use basic statistical functions: average, maximum, minimum, count, standard deviation.
- Copy the structure of a formula in one cell to another cell or a set of cells.
- Use the “Esc” key to remove the flashing dotted line around a cell, or to halt editing a cell and return to the original cell contents.

- Cut and paste, or Copy and Paste cell contents into other sections of the worksheet.
- Adjust column widths or row heights.
- Enable text entries to wrap around in a cell.
- Insert and delete rows and columns.
- Use the Chart Wizard tool to create charts
- Sort rows of data on one or more keys.

Most users who have already mastered the items on this list will also know a little about Pivot tables. Some of these items are covered in the following pages.

Select a range of cells, by “click and drag” or by using the Shift and “ Arrow” keys.

Most people who have had even a little exposure to spreadsheets know how to “click and drag” – unfortunately, this appealingly convenient method of “selecting” a range of cells is often the least effective and efficient, particularly when the range extends further than the screen boundaries. There are many ways of selecting larger ranges in Excel, but we will cover only a couple of ways that use keyboard shortcuts to select larger, contiguous blocks of data.

Below is a table of shortcut keys for selecting ranges.

Key	Extend selection from active cell to …
F8	Next cell selected
Shift-Arrow	Next cell selected
Shift-Home	Beginning of row
Ctrll-Shift-Home	Beginning of worksheet (A1)
Ctrl-Shift-End	End of worksheet (Most remote row, most remote column)
Shift-Space bar	Entire column of active cell
Ctrl-Shift-Space bar	Entire worksheet
Shift-PgUp	Cell in same column, one window up
Shift-PgDn	Cell in same column, one window down
Ctr-Shift-Arrow	Edge of the next block of data in the direction of the arrow key. Ctrl-Shift-right-down selects a rectangular block.
Shift-F5	Select one corner of the range; Press F5, enter the opposite corner of the range, and hold down the shift key as you choose the OK button.
Ctrl-Shift-*	Selects the smallest block which contains the current cell and all connected cells. “Connected” includes cells which touch at a corner.

Selecting multiple ranges
Excel allows you to select multiple non-contiguous ranges simultaneously. There are several ways, the simplest being:

1. Use the keyboard and Add mode: (ADD appears at the bottom of the screen)
2. Select the first range
3. Press Shift-F8 (to enter Add mode)
4. Select additional ranges
   (To leave ADD mode, press the Esc key)

Use the Mouse and the Ctrl key.

1. Select the first range.
2. Press Ctrl as you select each additional range.
3. Release the Ctrl key when finished.

This is handy sometimes when creating a data input area, for formatting a number of noncontiguous areas with a common format, or for charting. Multiple non-contiguous ranges

can be copied and pasted, provided they are aligned in the same rows or columns. When pasted, the gaps (rows or columns) disappear.

Enter text or numerical data into a cell.

Cells in Excel can contain text data, numerical data, or formulae and references. Entering data is generally fairly obvious. One additional feature in data entry that is sometimes useful is to create a data entry area. Below is an example of a data entry area set up in an Excel worksheet. As long as entries are separated by using the Tab key, the Enter key, or the Return key, the data entry area will remain selected (highlighted), and the cursor will remain within the data entry area. Pressing the Esc key, or one of the arrow keys will deselect the data entry area.

	A	B	C	D	E	F
1						
2						
3		Record 1		13	28	114
4		Record 2		15	33	110
5		Record 3		12	26	118
6		Record 4		12	27	
7		Record 5		14	34	

Notes:
1. The active cell is where the cursor is currently located, at cell F6
2. The data entry area is a noncontiguous area, selected using the Mouse and Ctrl key combinations.
3. The Enter key generally moves the cursor down after entry, while the Tab key moves the active cell to the right.
4. To reverse direction, modify the key with the Shift Key. That is, Shift-Enter moves up, while Shift-Tab moves left.
5. The active cell cycles through the data entry area – that is, after entering data into cell F7, the active cell will become B3.
6. When finished, pressing Esc, or clicking on any address with the Mouse will deselect the data entry area.

To correct an incorrect data entry, simply select the cell and type the correct entry over the original contents. Another alternative is to edit the contents, using either the command line, or double-clicking on the entry to enter the in-cell editor. Note that the navigational shortcuts all work within the cell editor: Home to the beginning of the line, End to the end, Ctrl-arrow to move left and right by a syntactic unit, and so on.

Move the contents of a cell into another location.

The easiest and most intuitive way to move a cell or range is to drag the cell or range to the new location and drop it. Under these circumstances, Excel moves the cell contents and the format. Alternatives may be to copy and paste, or to cut and paste. Copy, Cut, and Paste all appear as icons on the usual toolbar. However, it is usually faster to use keyboard shortcuts: Ctrl-X cuts, Ctrl-C copies, and Ctrl-V pastes. Note, however, that Ctrl-V pastes contents and format. A wider range of alternatives is offered by Paste Special, available from the Edit menu. (It is possible also to add a Paste Special icon to the toolbar).

Notes
Copy does exactly that – if the contents of the cell being copied is a formula, then the formula is pasted into the new location. All the relative addresses are moved the relative distances indicated by the size of the move; all absolute addresses, of course, are preserved under the Copy and Paste operation.
Cut is equivalent to a Move using a drag and drop operation: The original addresses are preserved, whether relative or not. This is sometimes a useful property when having to re-enter a long complicated formula to act on another part of the spreadsheet. The existing formula can be cut and pasted into the relevant relative address, as shown in the example below.
Example.
Below is shown the top part of a spreadsheet, including the first few rows and also the command line. The formula in cell L2 is shown in the command line.

Normal | Arial | 10

L2 = =IF(K2<10,1,IF(K2<50,2,IF(K2<100,3,IF(K2<500,4,IF(K2<1000,5,6)))))

	A	B	C	D	E	F	G	H	I	J	K	L
3867	96	0484 TC/2	TEACOSY	82.5	16.5	5	4	-1	-16.5	16.5	1	1
1	Category	Code	Description	On han	Cost Pr	Qty On	Stockta	Qty Var	$ Varia	AbsVar	AbsNum	NumCat
2	99	0115 NZFT	NZ FLAG TH	-39.6	1.2	-33	53	86	103.2	103.2	86	3
3	99	0048 PC001	PELLE SUED	273	6.5	42	1	-41	-266.5	266.5	41	2
4	84	0086 W/S8	WALKING S	179.49	59.83	3	1	-2	-119.66	119.66	2	1
5	101	0003 B98 3361	STEINLAGE	1822.5	67.5	27	24	-3	-202.5	202.5	3	1
6	56	0101 BS51	PURE SOUR	52.8	4.4	12	11	-1	-4.4	4.4	1	1
7	132	0058 KDTA	KEYRING O	71.5	6.5	11	0	-11	-71.5	71.5	11	2

This formula takes the value in cell K2, and transforms it into an integer between 1 and 6, according to a moderately complicated nested IF statement. The cell it operates on is immediately to its left, containing the numerical stock variance figure (as an absolute value), and all the addressing is relative. We wish to create a similar statement to categorise the absolute value off the dollar value of the stock variance figure, say DollarCat.

To use the same formula, reentering is tedious, but we need to be able to copy the formula in such a way that it refers to the AbsVar column. Two methods are immediately apparent.

1 Insert a column immediately to the right of the AbsVar column.
Copy the formula under NumCat into the cell now immediately to the right of the first AbsVar figure, in cell J2. The spreadsheet below is the result.

Arial 10

=IF(J2<10,1,IF(J2<50,2,IF(J2<100,3,IF(J2<500,4,IF(J2<1000,5,6)))))

B	C	D	E	F	G	H	I	J	K	L	M
484 TC/2	TEACOSY	82.5	16.5	5	4	-1	-16.5	16.5		1	1
:ode	**Description**	**On han**	**Cost Pr**	**Qty On**	**Stockta**	**Qty Var**	**$ Varia**	**AbsVar**		**AbsNum**	**NumCat**
115 NZFT	NZ FLAG TH	-39.6	1.2	-33	53	86	103.2	103.2	4	86	3
048 PC001	PELLE SUED	273	6.5	42	1	-41	-266.5	266.5		41	2
086 W/S8	WALKING S'	179.49	59.83	3	1	-2	-119.66	119.66		2	1
003 B98 3361	STEINLAGEI	1822.5	67.5	27	24	-3	-202.5	202.5		3	1

Note that the formula in cell K2 is exactly the formula we wanted. We can now enter a title above it, and duplicate the formula down the column. We might also want to move this column to a position alongside the NumCat variable. (Use drag and drop, or Cut and Paste)

2 Alternatively, we can Cut and Paste (or Move) the original formula one column to the right, as shown in the picture below:

Arial 10

=IF(K2<10,1,IF(K2<50,2,IF(K2<100,3,IF(K2<500,4,IF(K2<1000,5,6)))))

B	C	D	E	F	G	H	I	J	K	L	M
ode	**Description**	**On han**	**Cost Pr**	**Qty On**	**Stockta**	**Qty Var**	**$ Varia**	**AbsVar**	**AbsNum**		**NumCat**
.15 NZFT	NZ FLAG TH	-39.6	1.2	-33	53	86	103.2	103.2	86		3
148 PC001	PELLE SUED	273	6.5	42	1	-41	-266.5	266.5	41		2
186 W/S8	WALKING S'	179.49	59.83	3	1	-2	-119.66	119.66	2		1

Note the formula now in cell M2 refers to the cell two columns to its left. If we copy this and paste it into L2, we will have the appropriate formula for creating the DollarCat category. Again, placing a heading in L1, and then duplicating the formula down the column completes the task. The result is shown below:

C	D	E	F	G	H	I	J	K	L	M
Description	**On han**	**Cost Pr**	**Qty On**	**Stockta**	**Qty Var**	**$ Varia**	**AbsVar**	**AbsNum**	**DollarCat**	**NumCat**
NZ FLAG TH	-39.6	1.2	-33	53	86	103.2	103.2	86	4	3
PELLE SUED	273	6.5	42	1	-41	-266.5	266.5	41	4	2
WALKING S'	179.49	59.83	3	1	-2	-119.66	119.66	2	4	1

Final note: While the nested IF() statement is perfectly adequate for recoding these values to a 1-6 category, there are several other alternatives available for this, among them being the VLOOKUP() function, which might be a more attractive alternative here.

Erase the contents of a cell; Erase a whole set of cells at once.

To erase a range, including a single cell, or a multiple range, select the range and press the Del key. Note that pressing the BkSp key will erase the contents of a single cell, but will erase only the highlighted cell of a selected range. (Usually the top left cell)

You should try this in a spreadsheet, and remember to use the Undo command to restore the spreadsheet to its former condition. The keyboard shortcut for the Undo command is Ctrl-Z.

Use the fill handle to create a sequence of months, days, dates, times, or numbers.

You can use the fill commands on the edit menu to fill formulae or data into contiguous ranges. It is possible to fill Up, Down, Left or Right.

Here's what Excel's HELP file has to say about Fill:

Fill in a series of numbers, dates, or other items

1 Select the first cell in the range you want to fill, and enter the starting value for the series.

To increment the series by a specified amount, select the next cell in the range and enter the next item in the series. The difference between the two starting items determines the amount by which the series is incremented.

2 Select the cell or cells that contain the starting values.
3 Drag the fill handle over the range you want to fill.

To fill in increasing order, drag down or to the right.

To fill in decreasing order, drag up or to the left.

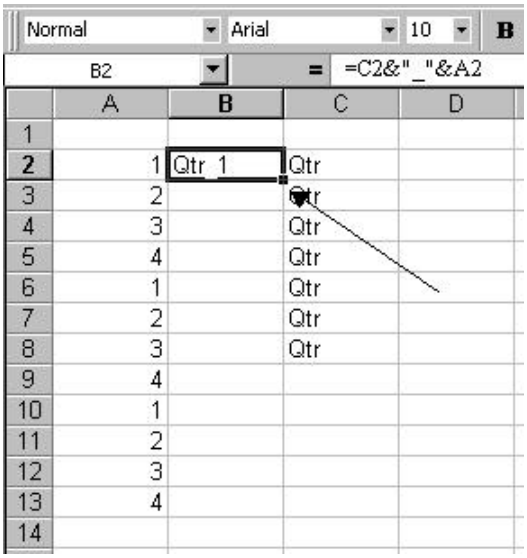
Tip To specify the type of series, hold down the right mouse button as you drag the fill handle over the range. Release the mouse button, and then click the appropriate command on the shortcut menu. For example, if the starting value is the date JAN-96, click Fill Months for the series FEB-96, MAR-96, and so on; or click Fill Years for the series JAN-96, JAN-97, and so on.

The Fill Handle is handy also for filling a row or column with a formula. The principle is the same: position the mouse pointer on the fill handle, and drag over the range you want to copy the formula through. Absolute and relative addresses in the formula behave exactly as you might suppose.

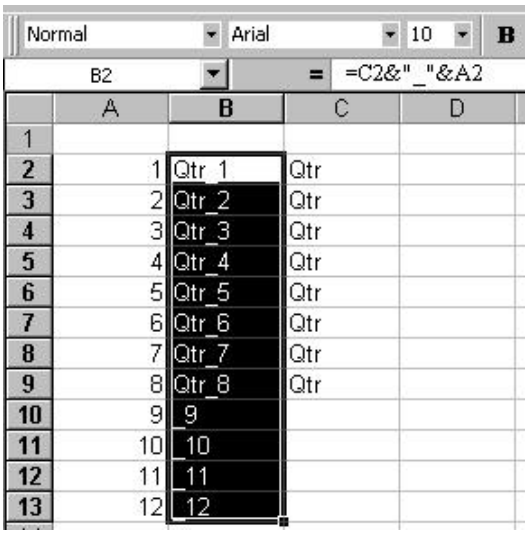
Double clicking.
Double clicking on the fill handle fills the formula down (not to the right) as far as the adjacent contiguous block of cells.
Example
Here is a fragment of a spreadsheet, with a formula entered into cell B2, as shown in the command line. The arrow in the picture identifies the fill handle.

Double-clicking on the fill handle in this situation will fill in Column B as far as cell B8, the end of the adjacent contiguous block. Notice that it does not fill to the furthest or nearest, but to the end of the contiguous block on the left.

Double-clicking on the fill handle to replicate a formula down a column is an enormously useful feature of Excel, particularly when working with large worksheets.



Below is the result of double-clicking on the fill handle.



Building formulae to add, subtract, multiply etc.
Use basic statistical functions: average, maximum, minimum, count, standard deviation.

Provided we have the correct formulae in place, it is generally wiser to let the computer do the work, rather than enter values that have been computed on a calculator. To illustrate the use of simple formulae, and how relative and absolute addresses can be used, consider a simple example, a two-way table of data.

Example
The data in the table below represent monthly tourist arrivals from Japan over a four-year period.

Inbound pax from Japan: 1993 - 1996				
	1993	1994	1995	1996
January	17108	18169	17230	19720
February	13613	15683	13370	17550
March	15329	16560	15800	18930
April	9733	11388	9830	12000
May	11419	12240	12090	12500
June	5839	6105	7970	7570
July	7616	6775	9060	8330
August	16318	18045	19970	19100
September	10756	10903	12200	12140
October	10469	11785	11100	11580
November	12098	13724	13870	13990
December	12320	15388	15340	14370

Our interest is finding the patterns in the table, using simple functions available in Excel.

What is the typical number of tourists arriving in New Zealand, from this table? This is simply the overall average, or arithmetic mean. The Excel function for calculating this is:

=AVERAGE()

All Excel functions have this form:

= <function name> (<list of arguments>)

Notice the opening and closing brackets. The <list of arguments> may be as simple as a single number, or may be several ranges within the worksheet. When Excel displays a formula, it displays all the function names in capital letters. This is quite handy: if you make a habit of always entering function names in lowercase, Excel will let you know when you mis-spell the function by continuing to display it in lowercase. Below is the fragment of spreadsheet above, with the average formula below and one cell to the right of the table.

VLOOKUP | =avrage(C4:F15)

	A	B	C	D	E	F	G	H
1								
2		Inbound pax from Japan: 1993 - 1996						
3			1993	1994	1995	1996		
4		January	17108	18169	17230	19720		
5		February	13613	15683	13370	17550		
6		March	15329	16560	15800	18930		
7		April	9733	11388	9830	12000		
8		May	11419	12240	12090	12500		
9		June	5839	6105	7970	7570		
10		July	7616	6775	9060	8330		
11		August	16318	18045	19970	19100		
12		September	10756	10903	12200	12140		
13		October	10469	11785	11100	11580		
14		November	12098	13724	13870	13990		
15		December	12320	15388	15340	14370		
16							=avrage(C4:F15)	
17								

(Note that the function has been mis-spelled.) The range to be averaged is rectangular, from C4 to F15. This the way that ranges are specified, as a cell address followed by a colon, then followed by a second cell address. Below is the result of entering the formula as shown in the display above. Note that an error message is displayed (#NAME?), and the function (shown in the command line) is shown in lowercase, demonstrating that Excel has failed to find it in its dictionary of functions and named ranges.

G16 | =avrage(C4:F15)

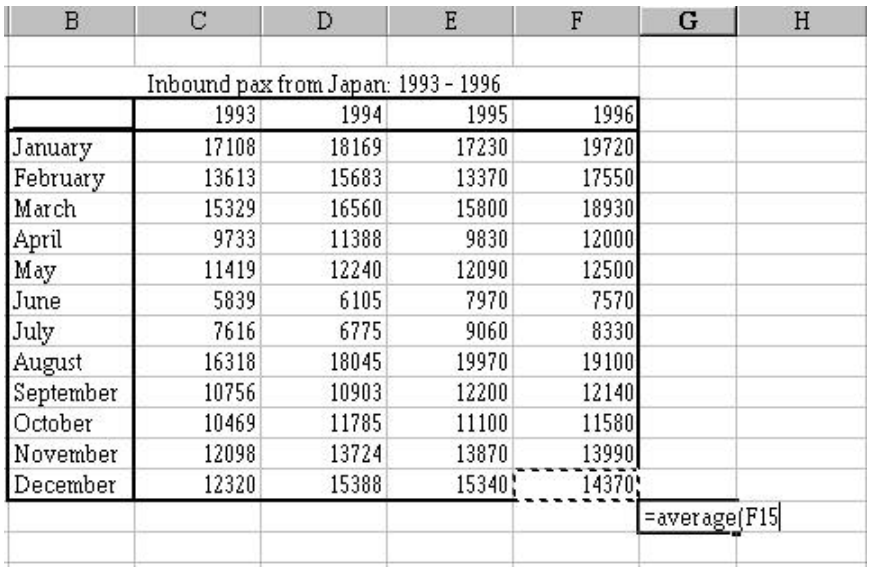
	A	B	C	D	E	F	G
1							
2		Inbound pax from Japan: 1993 - 1996					
3			1993	1994	1995	1996	
4		January	17108	18169	17230	19720	
5		February	13613	15683	13370	17550	
6		March	15329	16560	15800	18930	
7		April	9733	11388	9830	12000	
8		May	11419	12240	12090	12500	
9		June	5839	6105	7970	7570	
10		July	7616	6775	9060	8330	
11		August	16318	18045	19970	19100	
12		September	10756	10903	12200	12140	
13		October	10469	11785	11100	11580	
14		November	12098	13724	13870	13990	
15		December	12320	15388	15340	14370	
16							#NAME?
17							

The next display shows the result after the spelling mistake has been rectified. We now see that over this four-year period, an average of 13021 tourists has arrived in New Zealand from Japan each month.

Inbound pax from Japan: 1993 - 1996					
	1993	1994	1995	1996	
January	17108	18169	17230	19720	
February	13613	15683	13370	17550	
March	15329	16560	15800	18930	
April	9733	11388	9830	12000	
May	11419	12240	12090	12500	
June	5839	6105	7970	7570	
July	7616	6775	9060	8330	
August	16318	18045	19970	19100	
September	10756	10903	12200	12140	
October	10469	11785	11100	11580	
November	12098	13724	13870	13990	
December	12320	15388	15340	14370	
					13021

The easiest way to enter range addresses like this is to select the range, using either the mouse or the keyboard. The techniques discussed earlier, under "Select a range of cells, …" can be used here. In this example, following the steps below will create this formula.

1. In cell G16, type =average(
2. Press the left arrow, then the up arrow. This places the "marquee" or dotted outline, around cell F15. The screen looks like this:



B	C	D	E	F	G	H
	Inbound pax from Japan: 1993 - 1996					
	1993	1994	1995	1996		
January	17108	18169	17230	19720		
February	13613	15683	13370	17550		
March	15329	16560	15800	18930		
April	9733	11388	9830	12000		
May	11419	12240	12090	12500		
June	5839	6105	7970	7570		
July	7616	6775	9060	8330		
August	16318	18045	19970	19100		
September	10756	10903	12200	12140		
October	10469	11785	11100	11580		
November	12098	13724	13870	13990		
December	12320	15388	15340	14370		
					=average(F15	

3. Now hold the Ctrl key down, and the Shift key, and press the left arrow, then the up arrow. This places the marquee around the whole of the table. (Remember, Ctrl-Arrow moves over a contiguous block of cells. Holding down the Shift key continues to select.) The result of this selects both the left column of the table, where the months are entered, and the headings at the tops of the columns, where the years are entered. Since we don't want to include the top row and the left column, we want to "come back" one column, and one row.
4. Press the Shift key (and NOT the Ctrl key), and press the right arrow, then the down arrow. The result of this is to select the body of the table, including just the data we are interested in. Now enter a right bracket to create the formula for the average as required, as shown in the next display.

=average(C4:F15)

B	C	D	E	F	G	H
	Inbound pax from Japan: 1993 - 1996					
12R x 4C	1993	1994	1995	1996		
January	17108	18169	17230	19720		
February	13613	15683	13370	17550		
March	15329	16560	15800	18930		
April	9733	11388	9830	12000		
May	11419	12240	12090	12500		
June	5839	6105	7970	7570		
July	7616	6775	9060	8330		
August	16318	18045	19970	19100		
September	10756	10903	12200	12140		
October	10469	11785	11100	11580		
November	12098	13724	13870	13990		
December	12320	15388	15340	14370		
					=average(C4:F15)	

The value that results when the Enter key is pressed is 13021. Clearly, however, these data have significant monthly differences, and it appears that there is a year-to-year difference in the data. We will proceed with this example, and fit a model to the data to extract thyese yearly and monthly effects, to create a table of "expected" values using the fitted model.

In the same way that we have used the overall average to represent the "typical" value for the whole table, we can use the average of all the January values to get a typical value for January. Similarly, for all the other months, a typical value will simply be the appropriate arithmetic average. We will create a column of Monthly averages immediately to the right of the table, as follows:

1. In cell G3, enter a heading for the column, say Monthly Averages.
2. In cell G4, immediately below the heading, enter the formula for the average of the January figures. Try to use the keyboard method of selecting the range in the formula – practice makes perfect, and you <u>will</u> get faster at this!
3. When the averaging formula is entered, duplicate this formula down the column as far as December. The easiest way to do this is to use the Fill Handle – double-click on the Fill Handle, and the column will automatically fill down as far as the contiguous block of data on the left – i.e. as far as December. The display below shows the screen at the end of this step.

Tms Rmn 10

=AVERAGE(C4:F4)

	B	C	D	E	F	G
		Inbound pax from Japan: 1993 - 1996				
		1993	1994	1995	1996	Monthly Averages
	January	17108	18169	17230	19720	18056.75
	February	13613	15683	13370	17550	15054
	March	15329	16560	15800	18930	16654.75
	April	9733	11388	9830	12000	10737.75
	May	11419	12240	12090	12500	12062.25
	June	5839	6105	7970	7570	6871
	July	7616	6775	9060	8330	7945.25
	August	16318	18045	19970	19100	18358.25
	September	10756	10903	12200	12140	11499.75
	October	10469	11785	11100	11580	11233.5
	November	12098	13724	13870	13990	13420.5
	December	12320	15388	15340	14370	14354.5
						13021

Thus the January average is 18056.75, February is 15054 and so on. Similarly, we can extract column averages to get the 1993 average, the 1994 average, and so on. Follow a similar process (remember that you can't just double-click to Fill to the right!) to get the following display:

	A	B	C	D	E	F	G
1							
2		Inbound pax from Japan: 1993 - 1996					
3			1993	1994	1995	1996	Monthly Averages
4		January	17108	18169	17230	19720	18056.75
5		February	13613	15683	13370	17550	15054
6		March	15329	16560	15800	18930	16654.75
7		April	9733	11388	9830	12000	10737.75
8		May	11419	12240	12090	12500	12062.25
9		June	5839	6105	7970	7570	6871
10		July	7616	6775	9060	8330	7945.25
11		August	16318	18045	19970	19100	18358.25
12		September	10756	10903	12200	12140	11499.75
13		October	10469	11785	11100	11580	11233.5
14		November	12098	13724	13870	13990	13420.5
15		December	12320	15388	15340	14370	14354.5
16		Annual average	11884.8333	13063.75	13152.5	13981.6667	13021

We can think of these "typical values" as the best guess or estimate that we might be able to make with a limited amount of information. Thus our best guess for the monthly average inbound pax from Japan over this whole period is 13021. If, however, we restrict our attention to January, then our best guess would be 18056.75. If we were interested in "predicting" a value for June, we would use the June typical value, 6871. Similarly, if we wanted the best prediction for 1994, then the typical value for that year, 13063.75 would be

our best guess. We need to combine these two factors in some way to get a "prediction" for a given month in a given year.

We can approach this as a "predictor-corrector" type of problem. If our initial "model" is simply the overall average, how do we correct this value for a particular month? Clearly the result should correct the overall average 13021 back to whatever the appropriate row average is. The difference between the row average and the overall average is obviously the amount by which to correct the overall mean. Similarly, the difference between the column average and the overall average is the amount by which to correct for a given year. We can put these differences into our spreadsheet, using formulae that use a mixture of relative and absolute addressing.

Absolute addressing and relative addressing.
The normal reference that is written into a formula when using the mouse pointer, or the keyboard method, is a relative address: Copying the reference to a cell two cells to the right will change the reference to a cell two cells to the right of the previous reference and so on. To keep cell references from changing when you copy a formula to a new location (or use the Fill tool to propagate a function through a range of cells), use absolute references. A $ sign preceding either the row or the column reference (or both) will ensure that that particular part of the cell reference is unchanged under a copy operation. For a simple cell reference such as C17, the four possible combinations of absolute and relative references and their meanings are:

C17	Both references are relative. Both row and column will change by the relative amount of any move in the horizontal or vertical directions.
$C$17	Both references are absolute. Neither the row nor the column reference will change when the cell contents are copied to a new destination.
C$17	The column reference (C) will change by the relative amount when the formula is copied to another destination; the row reference ($17) will not change.
$C17	The row reference (17) will change by the relative amount when the formula is copied to another destination; the column reference ($C) will not change.

Returning to the exercise …..
Below is a display showing the entry in cell H4, giving the difference between the row average and the overall average.

	A	B	C	D	E	F	G	H
1								
2			Inbound pax from Japan: 1993 - 1996					
3			1993	1994	1995	1996	Monthly Averages	Row Effect
4		January	17108	18169	17230	19720	18056.75	=G4-G16

Now, as we replicate this formula down the column, we wish to successively calculate the differences between the rtow average and the overall average. Thus the formula in the next cell should be G5-G16, and so on. In other words, we want the G16 to be fixed – i.e. an absolute address. Thus the formula should be

=G4-$G$16

We can physically edit the formula, and enter the dollar signs, or we can use the F4 function. The F4 function key on the keyboard “toggles” the addressing in the address on which the cursor rests. Thus, with the cursor (in the command line) on the address G16, pressing the F4 function key will cycle through the four possible combinations of relative and absolute addressing. (And, on further pressing, will cycle through them again.) Pressing the F4 key oncewill yield the formula

=G4-$G$16

as desired. This formula can then be replicated down the column using the fill handle (etc) to produce this display:

	A	B	C	D	E	F	G	H
1								
2		Inbound pax from Japan: 1993 - 1996						
3			1993	1994	1995	1996	Monthly Averages	Row Effect
4		January	17108	18169	17230	19720	18056.75	5036
5		February	13613	15683	13370	17550	15054	2033
6		March	15329	16560	15800	18930	16654.75	3634
7		April	9733	11388	9830	12000	10737.75	-2283
8		May	11419	12240	12090	12500	12062.25	-958
9		June	5839	6105	7970	7570	6871	-6150
10		July	7616	6775	9060	8330	7945.25	-5075
11		August	16318	18045	19970	19100	18358.25	5338
12		September	10756	10903	12200	12140	11499.75	-1521
13		October	10469	11785	11100	11580	11233.5	-1787
14		November	12098	13724	13870	13990	13420.5	400
15		December	12320	15388	15340	14370	14354.5	1334
16		Annual average	11884.8333	13063.75	13152.5	13981.6667	13021	
17		Column Effect						

(Note that double-clicking on the Fill Handle will fill the column down as far as cell H16, which I have deleted.)

In a similar fashion we can produce a row of Column Effects, as the differences between the column means and the overall mean, as shown in the display below.

B	C	D	E	F	G	H
	Inbound pax from Japan: 1993 - 1996					
	1993	1994	1995	1996	Monthly Averages	Row Effect
January	17108	18169	17230	19720	18057	5036
February	13613	15683	13370	17550	15054	2033
March	15329	16560	15800	18930	16655	3634
April	9733	11388	9830	12000	10738	-2283
May	11419	12240	12090	12500	12062	-958
June	5839	6105	7970	7570	6871	-6150
July	7616	6775	9060	8330	7945	-5075
August	16318	18045	19970	19100	18358	5338
September	10756	10903	12200	12140	11500	-1521
October	10469	11785	11100	11580	11234	-1787
November	12098	13724	13870	13990	13421	400
December	12320	15388	15340	14370	14355	1334
Annual average	11885	13064	13153	13982	13021	
Column Effect	-1136	43	132	961		

The cells containing the averages and the "Effects" or correction figures, have been selected using one of the multiple range selection techniques discussed earlier, and formatted to appear as whole numbers.

The overall average (13021) together with the row and column correction factors or effects givge us a model for the data:

Data = Overall average + row effect + column effect

Thus this decomposition of the data gives us a "fitted value" for any combination of row and column (ie of month and year). The row effects summarise the monthly pattern, while the column effects summarise the annual pattern. For example, this model gives us, as a "fitted value" for 1994, July,

Fitted value = 13021 + (-5075) + 43

= 7989

(The actual value was 6775, so our model overestimates at this point.)

We'll now use Excel to create a table of fitted values from this model, illustrating the use of mixed absolute and relative addressing.

First, we will create an empty table with the same row and column headings as in the original table. The easiest way of doing this is simply to copy and paste the whole table, and then to delete its contents. The result of this is shown below, pasting the table into cell B20

	A	B	C	D	E	F
16		Annual average	11885	13064	13153	13982
17		Column Effect	-1136	43	132	961
18						
19						
20			1993	1994	1995	1996
21		January				
22		February				
23		March				
24		April				
25		May				
26		June				
27		July				
28		August				
29		September				
30		October				
31		November				
32		December				
33						

(The range from C21 to F32 is highlighted because this has just been selected and deleted.)

Now to fill the table with the fitted values, we will place an appropriate formula in cell C21, and replicate to the right (to cell F21) and then down (to cell F32).

In cell C21, type = (This begins formula entry)
Now click the mouse on the cell containing the overall mean, G16. As the formula is replicated right and down, we want this always to be G16, so press the F4 function key. This changes the entry to

=$G$16

Now enter +, so the cell entry is now

=$G$16+

and click the mouse on the first row effect, in cell H4. As the formula is replicated across and down, the column must remain the same, while the row must change according to whether it is January, February etc. Thus the column must be absolute, while the row must be relative. Press the F4 function key until the appropriate address is shown. (This should be three keypresses.) Now enter +, so the cell entry is

=$G$16+$H4+

and now click the mouse on the first of the column effects, in cell C17. In this case, the column must change, while the row must remain fixed, so the row is absolute ($17) while the column is relative. Two keypresses should accomplish this, so the final formuala is:

=$G$16+$H4+C$17

Pressing the Enter key will enter the formula and calculate the value.

Now place the cursor on cell C21, and, holding down the Shift key, press the right arrowthree times. This selects the range C21:F21. Ctrl-R will will this formula to the right. We can then double-click on the fill handle (Still with all of C21:F21 highlighted), and the formulae will fill down as far as the adjacent block of labels, the months. The resulting table is shown

below.

	A	B	C	D	E	F
16		Annual average	11885	13064	13153	13982
17		Column Effect	-1136	43	132	961
18						
19						
20			1993	1994	1995	1996
21		January				
22		February				
23		March				
24		April				
25		May				
26		June				
27		July				
28		August				
29		September				
30		October				
31		November				
32		December				
33						

Obviously we can now look at the residuals – the differences between the original data and our modeled values, and use these to give some idea of how good the fit is. This is reasonably straightforward, and graphical inspection of these values is a valuable method for exploring these data further. However, we have covered enough of absolute and relative addressing for now.

We have shown how to use the =AVERAGE() function, and how to use the absolute and relative addressing facilities of Excel. Other functions that might be useful in the present example include

=MIN()	Gives the smallest value in a range
=MAX()	Gives the greatest (the maximum) value in a range
=STDEV()	Gives the standard deviation of the values in the given range; this is a measure of the variability of the data in the range.

Clearly we can use the standard deviation function to give some idea of whether the row + column fit we have carried out is an improvement over simply using the overall average. The standard deviation of the original data was about 3700; the standard deviation of the residuals about the fitted values is approximately 750. This is a marked reduction in the error associated with our estimate of monthly pax, so this model represents an enormous increase in our ability to predict tourist numbers. (You may wish to pursue this part of the example as an exercise!)

Use the Chart Wizard tool to create charts

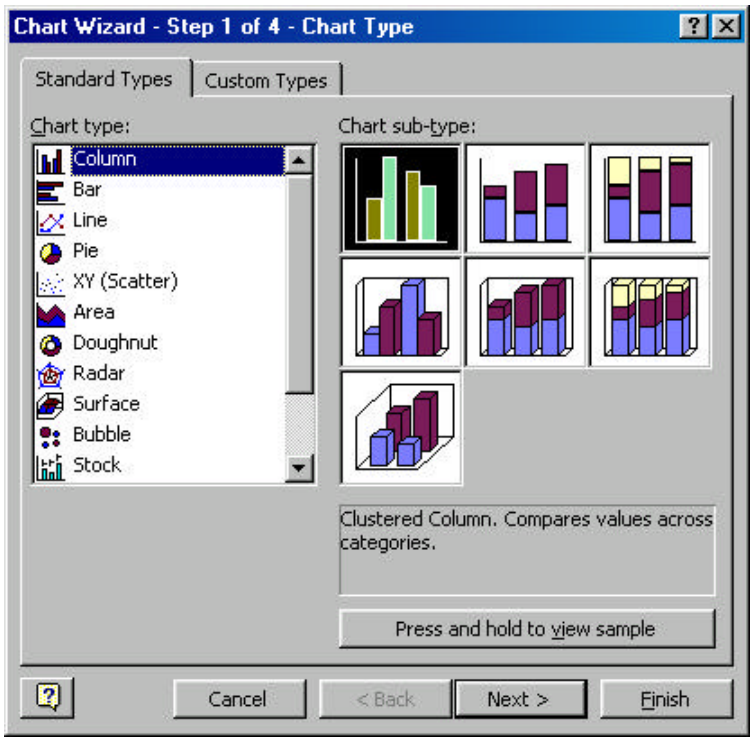
As with most things in Excel, there are many ways to create a chart from a set of data. We will illustrate the creation of a simple bar chart. Note that Excel has its own terminology for charts, not necessarily standard. For example, Excel differentiates between a horizontal bar chart and a vertical bar chart, calling the former a Bar Chart, and the latter a Column Chart. We will create a vertical bar chart (or Column Chart) from the data of the previous exercise.

Recall that we began with monthly inbound passengers from Japan into New Zealand, over the period from January 1993 to December 1996, and fitted a row + column model to the data. Among the estimates we calculated were the relative monthly effects, which we called the "Row Effects". We will create a chart of these row effects, to summarise the seasonal, or month-to-month effects observed in the data.

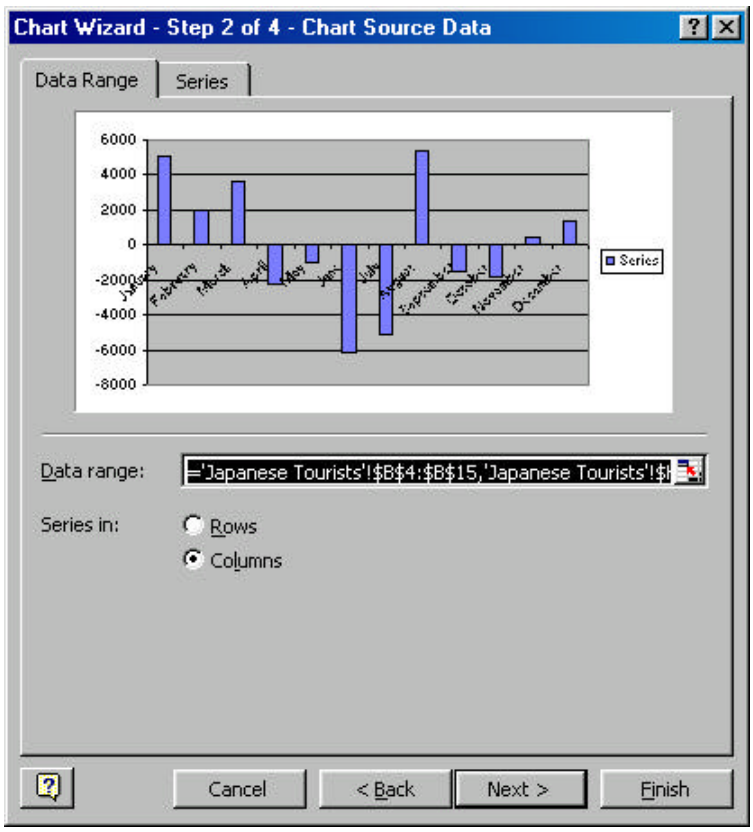
Below is the table of data, together with the row and column effects, as calculated earlier. The highlighted parts of the spreadsheet correspond to the months and their corresponding row effects. Note these areas of the spreadsheet are non-contiguous, and have been selected using the technique discussed earlier, using the Ctrl key to select a second, non-contiguous area.

B	C	D	E	F	G	H
	Inbound pax from Japan: 1993 - 1996					
	1993	1994	1995	1996	Monthly Averages	Row Effect
January	17108	18169	17230	19720	18057	5036
February	13613	15683	13370	17550	15054	2033
March	15329	16560	15800	18930	16655	3634
April	9733	11388	9830	12000	10738	-2283
May	11419	12240	12090	12500	12062	-958
June	5839	6105	7970	7570	6871	-6150
July	7616	6775	9060	8330	7945	-5075
August	16318	18045	19970	19100	18358	5338
September	10756	10903	12200	12140	11500	-1521
October	10469	11785	11100	11580	11234	-1787
November	12098	13724	13870	13990	13421	400
December	12320	15388	15340	14370	14355	1334
Annual average	11885	13064	13153	13982	13021	
Column Effect	-1136	43	132	961		3709

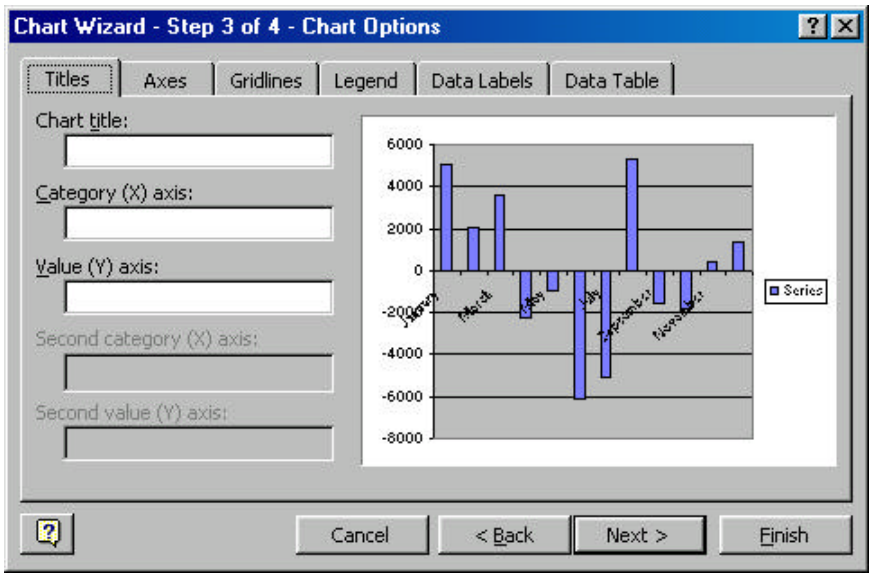
We can now click on the Chart Wizard icon on the menu bar. This results in the dialogue box below:



The chart type that we want is already selected, so we can simply press the "Next" button to carry on to the next dialogue box. Note the other types of chart available, and also note that the variety of charts extends beyond this display, as indicated by the "elevator: bar on the right of the list of chart types. Note also the tab with "Custom Types" on it – you can add a number of chart types of your own. Pressing Next gives us:

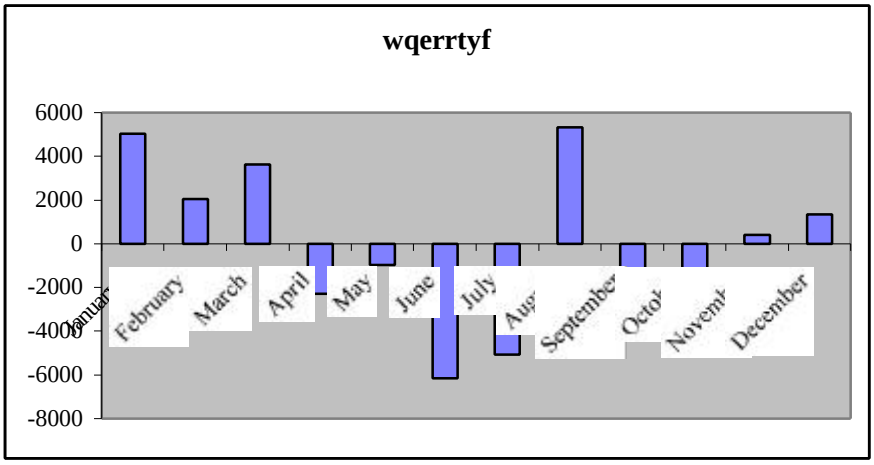


This gives us a preview of what the bar graph will look like. Pressing Next once more yields:



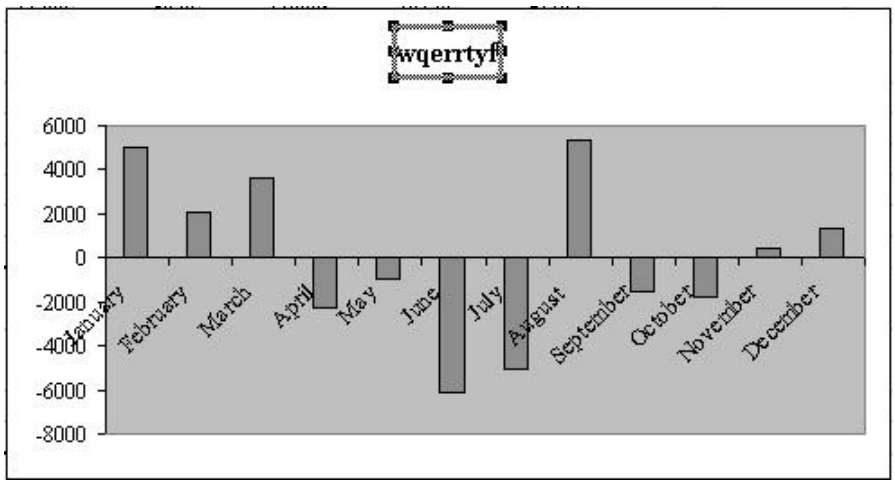
This dialogue box offers you the opportunity to customise the chart in a number of ways. The options available are: titles, axes, gridlines, legend, data labels, and data table. Each of these options has a number of component parts, enabling you to "tune" the appearance of the final

graph. I always put a dummy title on the chart (place the cursor in the Chart Title field and just strike a number of keys at random), remove the legend and the gridlines, and then accept the default values (for now!). The result of this in this case is shown below:



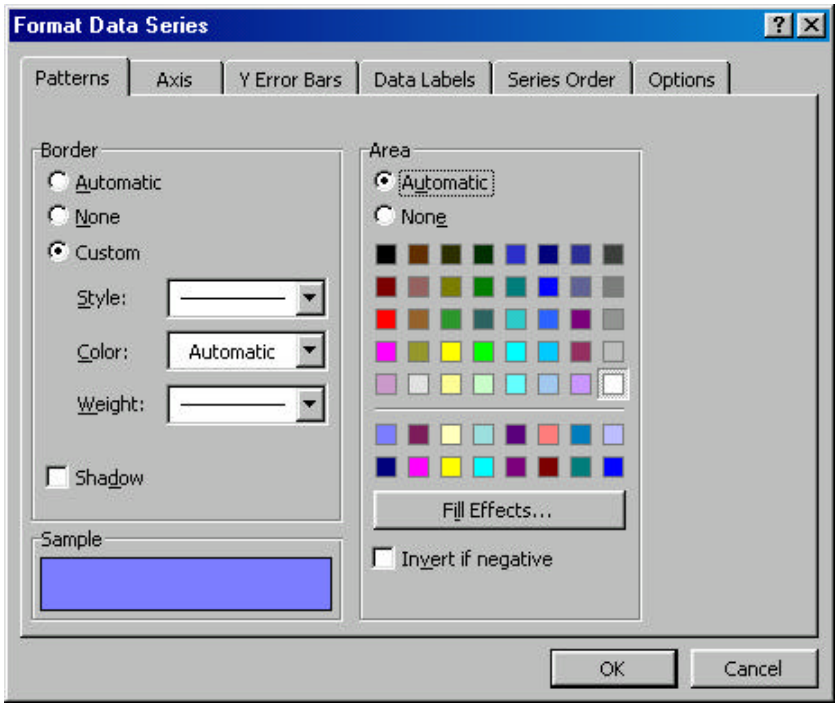
There are clearly a number of aspects to this chart that need to be addressed: the shading is not a particularly good idea for a chart to be printed in black and white; the column labels are very intrusive within the data plotting area; the title, of course, should be made meaningful; the chart should use the space that has been made available for charting. We will address a couple of these issues and try to "clean upp" the appearance of this chart.

First the title. My preference is for the title of a chart to be taken from the contents of a cell within the spreadsheet. Fortunately, this is very easily achieved. Simply click on the title (the chart then looks as below)

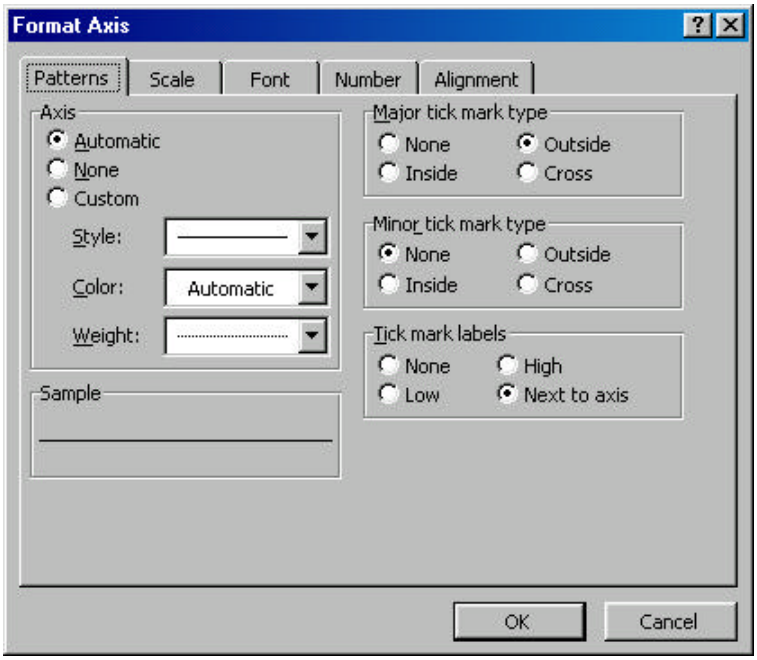


And then type = and point (within the spreadsheet) to the cell containing the text you want to use as the title. In this case, I entered the text "Seasonal effects in Japanese Tourists" into cell H2, clicked on the title as shown above, and typed = and then clicked on cell H2. The text then became the title of my chart.

Next, the shading. Having blue shading for the bars, against a blue background for the chart, doesn't seem very sensible. This is reasonably easily fixed. Double-click on the background of the chart. A palette of available colours comes up, from which we can select a background colour and also a colour for the border. I always select a white background and a black border, for maximum contrast. Now we can go through a similar process for the bars representing the data. The dialogue box for this is a little more complex: notice that the dialogue box also allows us to change colours and borders, as for the chart background, but also allows us to "tune" a number of the other characteristics of the chart. For now, we will content ourselves with the colours and shading. Again, to maximise contrast, I usually choose white with a black border.

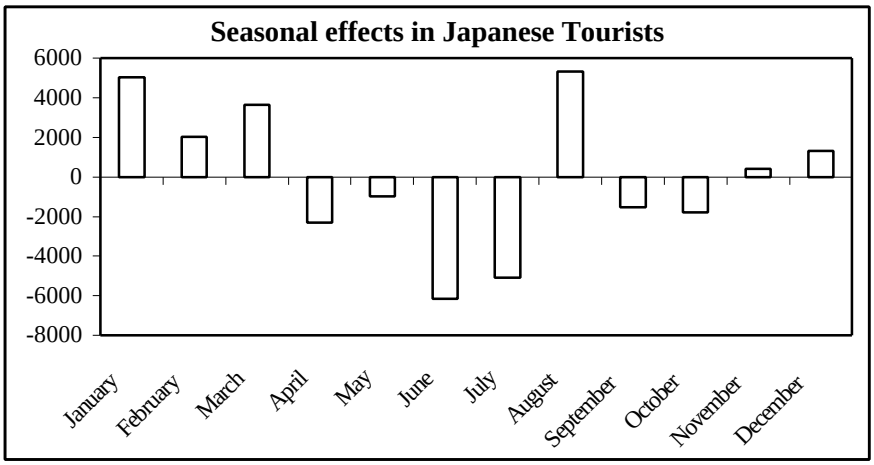


The labeling on the x-axis (the horizontal axis) intrudes into the plotting space and obscures,, rather than reveals, any patterns that might be evident in the data. We can edit a number of characteristics of the axis, as for the data series itself. If we double-click on the x-axis, this dialogue box results:



The key for us here is the field called "Tick mark labels". Within this array of choices, we want to select "Low", to place the labels along the bottom of the chart rather than immediately adjacent to the axis. We will allow the defaults on the other tabbed options within this dialogue.

Finally, we can change the shape of the chart itself within the box allocated to it by clicking on it, and then clicking on the handles to drag the chart into a different shape. The result of all these edits is shown below.



A final note regarding charts: the XY (Scatter) chart is the only "true" graphical tool in the Chart Wizard. Only the XY plot actually plots (x,y) pairs, while all the other chart types tend to have labels on the horizontal axis. If precise plots are required, it may be better to think of the requirement in terms of an XY plot, and create an appropriate template for this.

Pivot Tables

Pivot tables are an enormously powerful and useful feature of Excel, useful for summarising and exploring files that consist of "flat files" of a number of variables for a number of different observations. (Such layouts are called "lists" or "tables" in Excel.) Pivot tables are a type of crosstabulation of the data in a list. An example is shown below:

Sum of Quantity	Year													
Grape	88	89	90	91	92	93	94	95	96	97	98	99	(blank)	Grand Total
Burgundy							2		10					12
Cab Sauv							1		4	14	7			26
Cab Sauv										2				2
Cab Sauv - dessert									2					2
Chardonnay								1	8	31	45	9		94
Chateauneuf du Pape								1	3	5				9
Claret	1	2	13	3	2	6	15	21	62	50	35		2	212
Cotes du Rhone									2					2
Gewurzt/Riesling									2					2
Gewurztraminer				1	3	2			2	2				10
Merlot								5	7	1	5			18
Muscat				1										1
Pinot Gris										2	5	6		13
Pinot Noir										10	13	7		30
Pinotage										2				2
Riesling						1	7	1	13	4	12	6		44
Riesling-bot			3	1					3		3			10
Sauternes		1												1
Sauvignon Blanc				4	1									5
Semillon						1					1			2
Syrah							5	3	10	11	4			33
Grand Total	1	3	16	10	6	10	30	32	128	134	130	28	2	530

This pivot table gives a two-way classification of the bottles of wine in a friend's cellar. The classification of Grape (or wine) type by vintage gives a view of the coverage that the cellar offers, shows where the cellar has weaknesses, and also gives some indication of the owner's preferences! We will revisit a previous example to show how pivot tables can be used to investigate quite large datasets.

Example

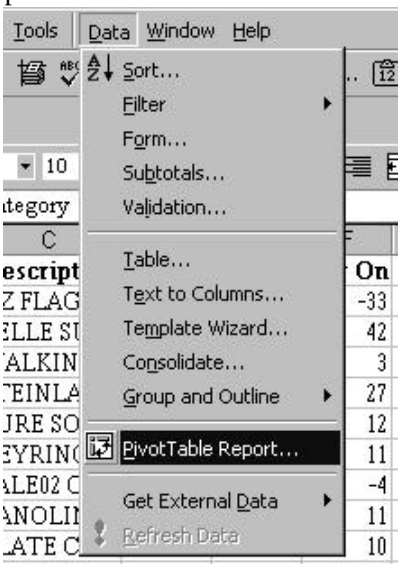
Recall the example we used in the section on moving the contents of a cell into another location. The data arose from a stocktaking exercise at a small retail operation, and resulted in some 3866 records of items that did not correspond to the records in the computer system. We looked at only the top few lines of data in the dataset, and added a couple of new fields, NumCat and DollarCat, as shown in the fragment below:

C	D	E	F	G	H	I	J	K	L	M
Description	On han	Cost Pr	Qty On	Stockta	Qty Var	$ Varia	AbsVar	AbsNum	DollarCat	NumCat
NZ FLAG TH	-39.6	1.2	-33	53	86	103.2	103.2	86	4	3
PELLE SUED	273	6.5	42	1	-41	-266.5	266.5	41	4	2
WALKING S	179.49	59.83	3	1	-2	-119.66	119.66	2	4	1

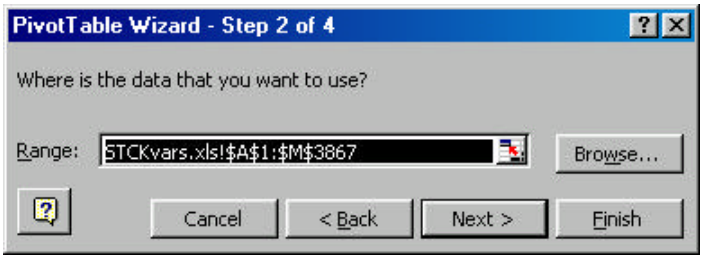
We will now form a Pivot Table with these data, and use the NumCat and DollarCat categories as row and column classifications, and in the body of the table, we will form the sum of the absolute values of the dollar variances.

Step 1

Select the Pivot Table report option from the Data menu:

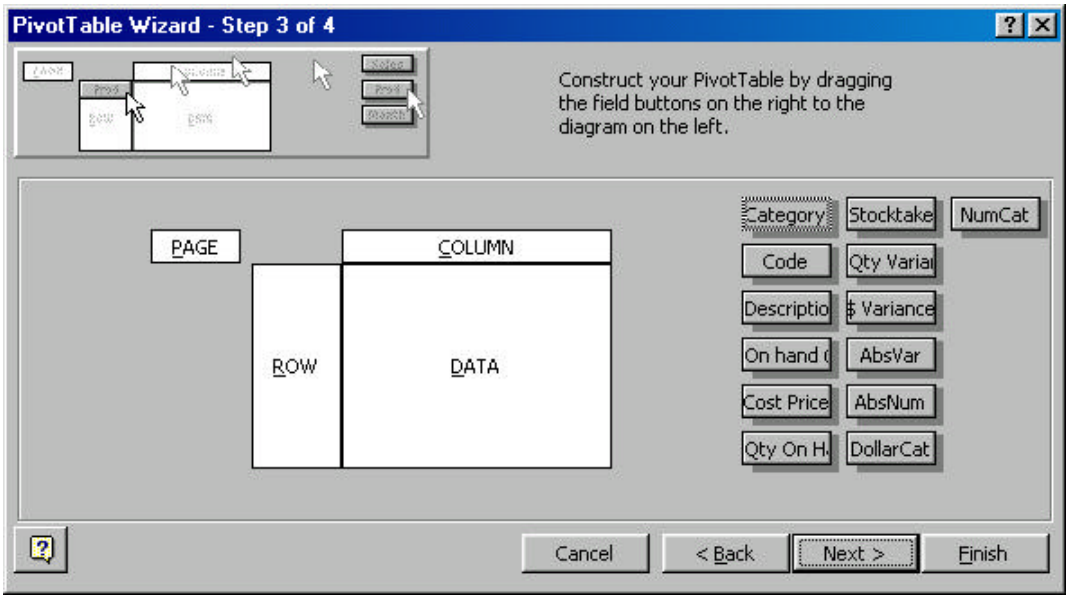


Step 2
The data are in an Excel list. Notice that the Pivot Table tool can take data from other sources. Click on the Next button.

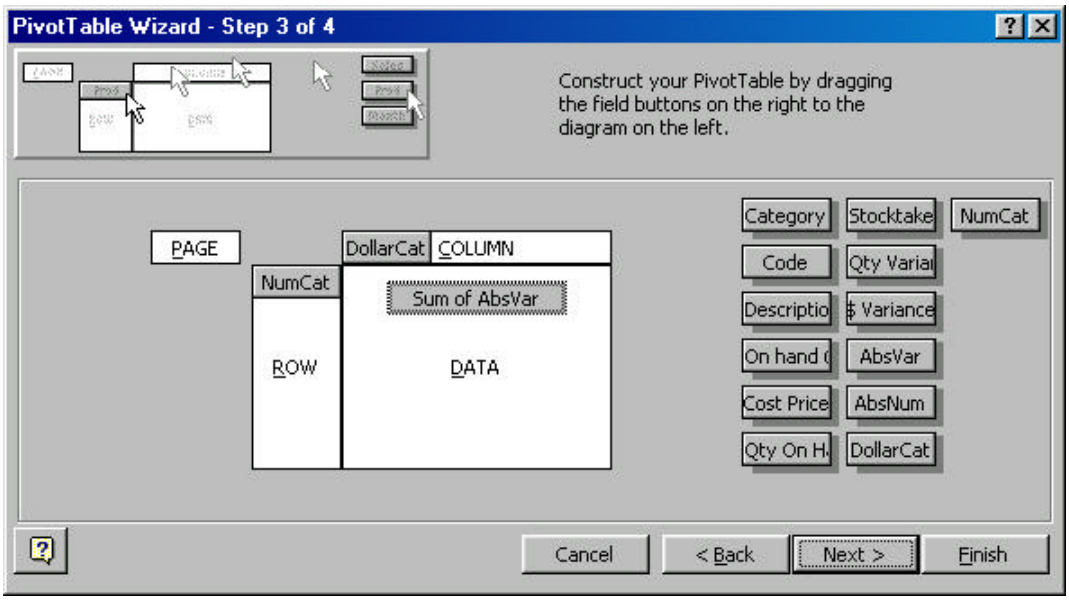


If the range given in this dialogue box does not cover the required range, change it. Note that the range should include titles for the columns of data. These titles will be used as field names for the construction of the Pivot Table.

Step 3
The next dialogue box is used to lay out the Pivot Table. (Versions of Excel later than Excel 97 have a slightly different set of dialogue boxes that accomplish the same thing.) Each of the variables in the list is called a field, and the fields (here, Category, Code, Description etc) are laid out in boxes, ready to be dragged into position on the Pivot Table. In our case, we will drag NumCat into position as the row classification, and DollarCat into position as the column classification. The data field to be used for the interior of the Pivot Table will be AbsVar, the absolute value of the dollar variance. Notice that in the Pivot Table layout, we chosen the Sum of the absolute values of the dollar variances as the data field to display in the table. This can be edited by double-clicking on the field, and choosing whatever summary statistic is most appropriate. The bare table, and the result of choosing the fields for analysis, are both displayed below.



and after laying out the table, we have:



Step 4

If we now click on the Next button, we get the dialogue box below, inviting us to place the resulting Pivot Table in a sheet of its own, or in the current sheet at a specified address. The default is for a new sheet, so we will run with this for now. The resulting Pivot Table is (Note that I have reformatted all the entries to be integers):

Sum of AbsVar	DollarCat						
NumCat	1	2	3	4	5	6	Grand Total
1	5336	34331	43430	82399	31327	31484	228307
2	74	2809	3992	15393	6615	1455	30337
3		89	348	1600	3377	3687	9101
4	0	47	63	1861	731	1254	3955
5					563		563
6	0			709	1816	8869	11394
Grand Total	5409	37276	47833	101961	44429	46749	283657

Thus the cell containing $31484 in absolute terms, at the top right corner of the table, corresponds to items with a high dollar value (recall that DollarCat=6 means that the variance was in excess of $1000), and quite low frequency (recall that NumCat=1 means that there were fewer than 10 items accounting for the variance). A different view of the data can be obtained by changing the basis on which the Pivot Table is calculated. This is easily done – we can edit the Pivot Table by placing the mouse pointer on the table and right-clicking. The dialogue box is shown below:

Choosing the Wizard option allows you to change the layout, including the variables used to construct the Pivot Table. If we change the layout to give us the "Sum of AbsNum" as the tabulated field, then this Pivot Table results:

Sum of AbsNum	DollarCat						
NumCat	1	2	3	4	5	6	Grand Total
1	1698	3153	1195	962	142	12	7162
2	270	1574	1180	1638	156	15	4833
3		120	328	532	354	146	1480
4	538	117	114	774	261	165	1969
5					662		662
6	9342			2402	4160	9336	25240
Grand Total	11848	4964	2817	6308	5735	9674	41346

Thus we can now see that there are only 12 items that contribute to the $31484 total variance in the first table. The power of the Pivot Table begins to become apparent! A further characteristic that makes Pivot Tables enormously useful is that we can double-click on an entry to "drill down" to see exactly what the contributing records are from the original list. If

we double-click on the 12 in the 1-6 cell, we get these records:

Category	Code	Description	On hand @ W.av cost	Cost Price	Qty On Hand	Stocktake Actual	Qty Variance	$ Variance	AbsVar	AbsNum	DollarCat	NumCat
132	0532 TEESI S	Pearl-BLAC	1400	1400	1	0	-1	-1400	1400	1	6	1
154	0020 JE253	Pearl-BLAC	3000	3000	1	0	-1	-3000	3000	1	6	1
156	0870 BP22	Pearl-BLAC	6000	6000	1	0	-1	-6000	6000	1	6	1
152	0020 ZZ PO863	Pearl-BLAC	3000	3000	1	0	-1	-3000	3000	1	6	1
167	0122 GB 003 S	BLACK EARR	1012	1012	1	0	-1	-1012	1012	1	6	1
152	0207 HP14 P	WHITE CAB	2072	2072	1	0	-1	-2072	2072	1	6	1
122	0004 SR1426	Pearl-SEMI	1800	1800	1	0	-1	-1800	1800	1	6	1
154	0401 J77 JC	15.42CT WH	1400	1400	1	0	-1	-1400	1400	1	6	1
149	0410 E3546	CRYSTAL CA	3200	3200	1	0	-1	-3200	3200	1	6	1
150	0410 RO267118C	BLACK CAB	5800	5800	1	0	-1	-5800	5800	1	6	1
132	0532 TEEHA S	Pearl-SEMI	1300	1300	1	0	-1	-1300	1300	1	6	1
132	0532 TEEHA L	Pearl-BLAC	1500	1500	1	0	-1	-1500	1500	1	6	1

The clustering of these items together is the sort of valuable information that makes stocktaking exercises worthwhile! There is obviously a great deal more useful information that we could extract from this Pivot Table.

There are many other valuable properties of Pivot Tables, and ways in which to use them: a great deal of information is available in the Help files, and in reference books. Note also that we could have been much more clever in terms of labeling the rows and columns of the table, to make them more readable.