

Excel 2002

Objectives

This unit builds on the previous *Quick Start to Excel* by introducing the more advanced concepts which are applicable to using Excel in a commercial context.

The objective of this unit is to enable you to:

- Make use of the data entry tools: AutoFill, AutoComplete, AutoCorrect, Undo.
- Apply advanced cell formatting: text alignment, AutoFormat, automatic formatting, conditional formatting and more advanced number formats.
- Use the following built-in functions: sum, average, count, max, min, var, stdv, median, round, int, abs, ceiling, floor, rand, choose, vlookup, hlookup, now.
- Know how and when to use absolute, relative and mixed cell references.
- Assign and use range names.
- Understand the basics of Excel's date and time arithmetic.
- Know how to work with larger spreadsheets, including window zoom, full screen mode, collapse buttons and split windows.
- Use the values in a spreadsheet to generate a chart and customize the chart.
- Use the grouping feature and data filter feature to assist the viewing of larger data sets.
- Use pivot tables for simple data mining.
- Describe how to construct a good spreadsheet.



Efficient Data Entry

You will spend quite a bit of time entering data in the spreadsheet. Excel has tried to make things as easy as possible for you by including “intelligent” data entry tools: AutoFill, Auto-Complete, AutoFormat and AutoCorrect.

Using AutoFill

Copying the contents of a cell or range to adjacent cells.

You can use the clipboard copy and paste functions from the Edit menu or the standard toolbar to copy data from one cell or range to any other cell or range. However, if you want to copy (or repeat) the contents of a cell into adjacent cells, follow these steps.

1. Click the cell (to make it the current cell). Alternatively, you can select an entire range if you wish to copy the contents of more than one cell.
2. Move the mouse cursor over the *handle* on the bottom right corner. The cursor will turn from a white cross  into black cross .
3. Drag the cursor vertically *or* horizontally along as many adjacent cells as you wish to copy the data to.

Using AutoFill for common data series.

Excel makes intelligent use of the above procedure to save you from the tedious job of having to type in frequently used data series.

- If a cell contains the text **Mon**, dragging the cell handle will automatically fill the adjacent cells with Tue, Wed, Thu, Fri etc. Similarly, **Monday** will be “extended” into Tuesday, Wednesday, Thursday etc.
- If a cell contains the text **Jan** (or **January**), AutoFill will complete the sequence of the months Feb, Mar, Apr ... for you if you drag the cell handle across adjacent cells.
- Similarly, **1st** will be extended into 2nd, 3rd etc.
- AutoFill will even handle an entry like **Product1** and modify it to become Product2, Product3 etc.
- If a cell contains any other value that Excel doesn't recognise as part of a standard sequence, it will just copy the value (label or number).

Note: if you wish to create a number series which increases by **1** from cell to cell, you can use a modified version of the AutoFill command: just hold down the **<Ctrl>**-key while dragging the cell handle.

Using AutoFill to enter a linear series of data.

AutoFill can also be used to enter a series of (linearly) increasing or decreasing data. Assume that you wish to enter the following series in a column: 2, 4, 6, 8, 10 ... up to 50. Use *AutoFill* in the following way to make your work a lot easier.

1. Enter the number **2** in the first cell of the column.
2. Enter the number **4** in the second cell of the column.
3. Highlight (select) both these first two cells of the column with your mouse cursor.
4. Move the mouse cursor to the (bottom right) cell (range) handle and drag it vertically down the column. As you drag the handle across the cells, you will see AutoFill's suggested values appearing: 6, 8, 10 etc. Keep dragging down until you reach your desired final value (e.g. 50).

	A	B	C	D	E	F	G	H	I
1	Some examples of AutoFill								
2									
3	You typed:	Mon	Thursday	March	1st	11th Feb	Exercise 1	3	09:00
4	AutoFill can complete these entries for you:	Tue	Friday	April	2nd	12th Feb	Exercise 2	3	10:00
5		Wed	Saturday	May	3rd	13th Feb	Exercise 3	3	11:00
6		Thu	Sunday	June	4th	14th Feb	Exercise 4	3	12:00
7		Fri	Monday	July	5th	15th Feb	Exercise 5	3	13:00
8		Sat	Tuesday	August	6th	16th Feb	Exercise 6	3	14:00
9									
10	You typed these 2 cells:	1	Mon	Dec	1999	05-Mar-98	1st Term	10	09:00
11		2	Wed	Oct	1998	12-Mar-98	Totals	7.2	09:15
12	AutoFill can complete these entries for you:	3	Fri	Aug	1997	19-Mar-98	2nd Term	4.4	09:30
13		4	Sun	Jun	1996	26-Mar-98	Totals	1.6	09:45
14		5	Tue	Apr	1995	02-Apr-98	3rd Term	-1.2	10:00
15		6	Thu	Feb	1994	09-Apr-98	Totals	-4	10:15
16		7	Sat	Dec	1993	16-Apr-98	4th Term	-6.8	10:30

AutoFill can handle a wide variety of series, including dates. The following are some examples of series that are completed if you enter their first two (cell) values.

- 3-Mar, 10-Mar è 17-Mar, 24-Mar, 31-Mar, 7-Apr ...
- Jan, Mar è May, Jul, Sep, Nov ...
- 0, 10 è 20, 30, 40 ...
- 6, 4 è 2, 0, -2, -4, -6 ...
- Product 1, Order è Product 2, Order, Product 3, Order, Product 4, ...

AutoFill can be used for even more advanced series. Check *Excel Help* for detailed instructions on how to use the following features.

- It can extrapolate a linear trend from numbers that are not exactly evenly spaced:
e.g. 1, 2, 4 è 5.33333, 6.83333, 8.33333, 9.83333 ...

- It can be configured to complete series by multiplying with a constant (instead of adding a constant): e.g. 1, 2, 4 è 8, 16, 32 , 64 ...
- You can define your own frequently used custom series
e.g. 2, 3, 5, 7, 11 è 13, 17, 23, 29, 31 ...

AutoComplete

While you are typing a cell entry, *Excel* sometimes even guesses what you are going to type next based on what you have typed so far. When you are typing an entry that starts the same way as other entries in the same column (or row), it will suggest to complete your entry with the remaining characters of that entry. E.g. when you are entering a series of products such as margarine, butter, cheese ... in consecutive cells and you type “m” in a next cell, *Excel* will suggest the completion (m)“*argarine*” for you. You can press **<Enter>** to accept the suggested completion or just continue typing if you need to enter some different data.

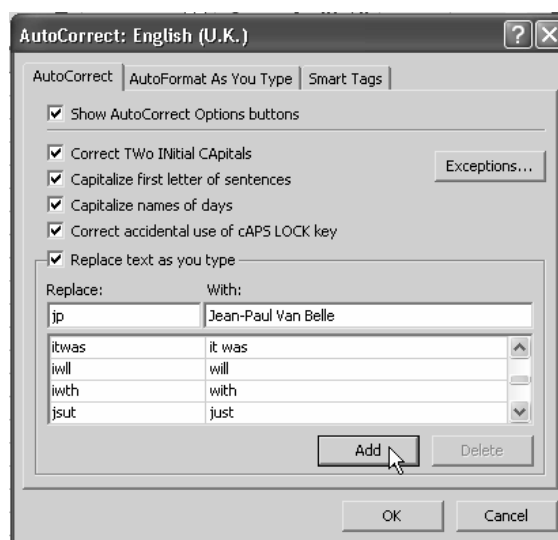
	A	B
1	Product	Kg sold
2	margarine	20
3	butter	10
4	cheese	5
5	margarine	

The difference between *AutoComplete* and *AutoFill* is that *AutoComplete* automatically suggests a completion for text entered within the same cell; *AutoFill* fills other (empty) cells based on the contents of previous cells.

Other Automatic Data Entry Tools

Like MS-Work, *Excel* offers many other tools that can speed up data entry.

AutoCorrect : will automatically correct typing mistakes, replace abbreviations with full length versions and insert common typographic symbols e.g. typing (c) will automatically be replace with ©. I have added an AutoCorrect entry (in *MS-Word!*) which substitutes my initials **jp** with **Jean-Paul Van Belle**.



Spell check : you can spell check a cell entry, a range or your entire spreadsheet by pressing the **<F7>** function key (or select the **Spelling...** option from the **Tools** menu).

Undo : if you realise that what you have typed or done is incorrect, press the **<Ctrl>-<Z>** “undo” key (or click the *Undo* button on the Standard Toolbar).

Find and Replace : a quick way of changing data entries across the entire spreadsheet (or a given range) is by selection the **Replace** option from the **Edit** menu (or press **<Ctrl>-<H>**).

AutoFormat : recognises and applies the appropriate display format when entering a number e.g. a date, currency or percent. Described under *More Advanced Cell Formatting*.

More Advanced Cell Formatting

The following screen gives an overview of the various types of formatting which can be applied to spreadsheet cells.

	A	B	C	D	E	F
1	SA Sausage Sorcerer					
2	Sales (1st Quarter)					
3		Jan	Feb	Mar	Total sales	%
4	Boerewors	R 3,400	R 3,210	R 3,385	R 9,995	37%
5	Morreensburg Special	R 2,145	R 2,250	R 2,640	R 7,035	26%
6	Pofadder Game Sausage	R 1,200	R 1,890	R 1,420	R 4,510	17%
7	Willy's Wors	R 2,120	R 1,300	R 1,965	R 5,385	20%
8	Total sales:	R 8,865	R 8,650	R 9,410	R 26,925	100%

The Formatting Toolbar

The most frequently used formatting options are available from the standard Office *formatting toolbar*.




The options include the following (from left to right)

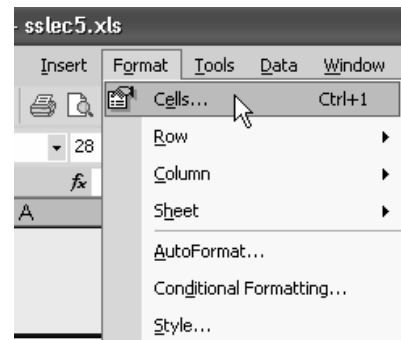
- **Font type:** the selection of *Windows* fonts depends on your specific installation, but will usually include Arial, Times (New) Roman, Courier, (Brush) Script and many fancier ones.
- **Font size:** select a default size from the drop-down menu or enter any size of your liking (size is measured in points, 30 points measure 1 centimetre when printed)
- **Font style:** apply any combination of the standard styles: bold, italics and/or underline.
- **Cell alignment:** align cell entries to the left, centred or right aligned. Alternatively, information can be centred and merged across several (previously selected) cells.
- **Number format:** select the Currency, Percentage or Comma format and increase/decrease the number of decimal places to be displayed.
- **Indent:** indent information by one or several tabs.
- **Borders:** select a cell border style and width from the drop-down menu.
- **Colours:** select a cell background ("*fill*" or highlight) or text colour.

All of the above options were discussed in the *Easy Start to Excel*; you should have practised the formatting toolbar buttons on simpler spreadsheets.

Selecting more advanced cell formatting

There are a number of different ways in which you can apply cell formats that are not available from the formatting toolbar. The following are the different ways in which to open the *Format Cells* window.

- From the *Format* menu (click on the menu option or press **<Alt>-<O>**) select the **Cells...** option (see right).
- Right-click a cell and select the **Format Cells...** option  from the pop-up menu.
- Press the **<Ctrl>-<1>** shortcut key combination to open the *Format Cells* window directly.



If you find that you are using a particular format very frequently, consider adding it to the formatting toolbar using the *Tools • Customize... • Toolbars* menu option. Alternatively, create your own shortcut key for your favourite format option.

Cell formatting can be applied in two ways.

1. You can first select a format and then you type text – all information that is typed in the *current cell* will inherit this format.
2. The other (usual) way is to first enter the text (which will inherit the default cell format) and afterwards select (a portion of) the text and apply the format to the highlighted text.

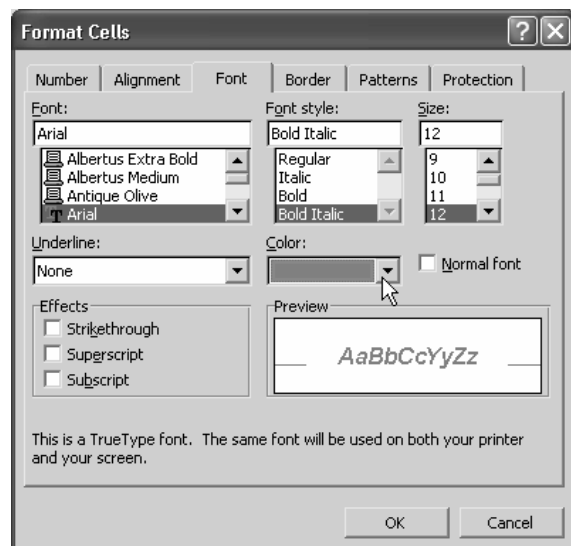
Text within one single cell can have different *font* types, styles and sizes. If you select the *Cell Format* command while you are entering text in a cell, only the *font* tab will be available.

Font Formats

Click the *Font* tab in the *Format Cells* window to access the font options. Most options are available from the formatting toolbar, except for the following:

Underline style: you can select double or accounting underlining; the latter is slightly “lower” than standard underline.

Effects: to select Superscript/ Subscript and Strikethrough effects.

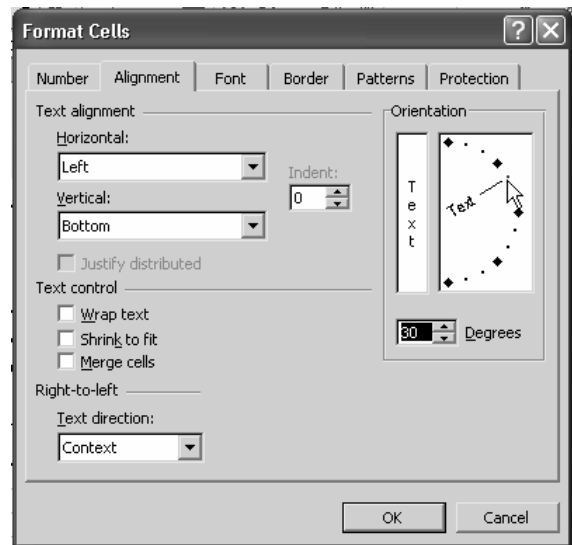


Text Alignment within Cells

The alignment tab allows you to control the position of text (or numbers) with respect to the boundaries of the cells. You have three generic control options.

Text Alignment.

- The **Horizontal** alignment determines the position in respect of left and right borders of the cell. A special option worth mentioning is the ***Fill*** option, which repeats the text until it fills the entire cell.
- The **Vertical** alignment determines the vertical position of the text. Text can be positioned at the top, in the middle or at the bottom of the cell.
- The **Indent** option allows you to indent text from the cell borders.



Text Control

This allows you to specify what must happen should the length of the text exceed the space available in the cell.

The ***Shrink to Fit*** option will reduce the size of the text until the text can be displayed in the space available, whereas the ***Wrap Text*** option will wrap the text over several lines. Wrapping text will increase row height and, unless row height is set to *Auto*, you may need to adjust the row height manually.

Text Orientation

Use this option to rotate text from +90 (vertical, text reads from bottom to top) to -90 degrees. Drag the red handle at the end of the line or enter the desired number of degrees.

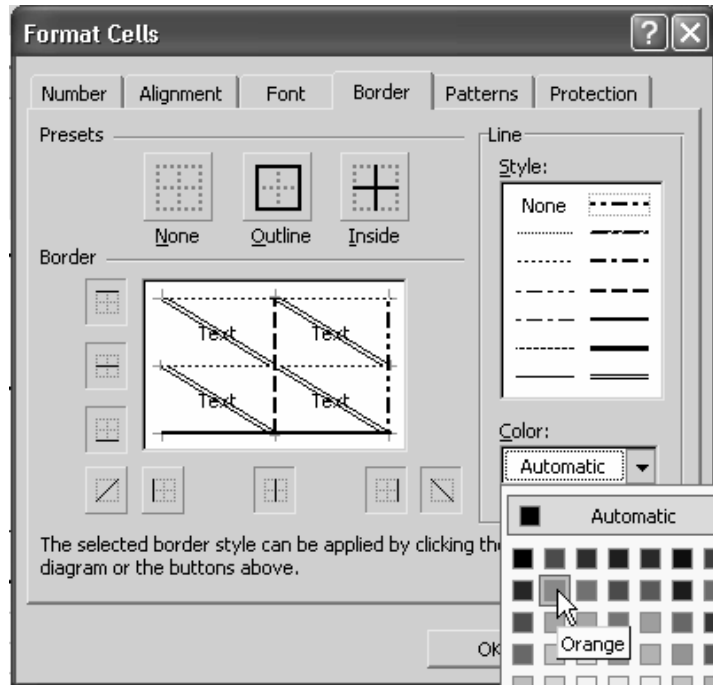
	A	B	C	D
1	Horizontal alignment	Other	Vertical alignment	Orientation
2	Left aligned	This text wraps around but is not justified.	Text at top	Vertical or +90 degrees
3	Centred	This text wraps around and is justified.	Text middle	45 degrees
4	Right aligned	Text is set to <i>Shrink to Fit</i>	Text at bottom	30 degrees
5	Fill align.Fill align.Fill align.			
6	This text doesn't wrap but "spills over" into adjacent empty cells.			
7	This text is centred across four cells/columns.			

Borders, Colours, Shading and Patterns

The use of cell borders, colours, background shading and patterns can enhance the appearance of text in a cell.

Although **borders** around a cell can be created quickest by using the formatting toolbar icon, you have much more control when using the *Border* tab of the *Format Cells* window. This window allows you to select different line styles, even combining them within the same cell!

If you have selected a *range* of cells, you can apply the borders to the *Outline* (i.e. the outside) of the range or between the cell boundaries of the cells making up the range i.e. the *inside* (or *delete* borders by selecting the *None* preset).



You can also select a specific *colour* for your border or even apply *diagonal* lines *through* your cell.

Click on the **Patterns** tab to have a wider variety of background colours for your cells than is available from the *formatting toolbar* button. This will also allow you to select from a variety of background patterns.


The foreground colour (i.e. colour of the text) can be set by clicking the *Font Colour* button on the *formatting toolbar* or by using the *Font* tab of the cell formatting window.

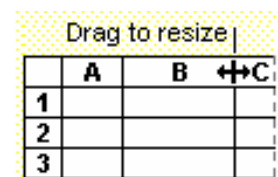
Row and Column Formatting

In addition to cell formatting, there are some additional formatting options that apply to columns and rows only.

Changing the width of a column

You can change the width of (all the cells in) a column in one of the following ways:

- Click the *Column* option from the *Format* menu and select  *Width...* to specify the exact desired width (in characters).
- Drag the *right* border of the *Column Name* (or column heading) until you have the desired width.
- Double-click the *right* border of the *Column Name* to set the



column width automatically to fit the longest text entry in the column.

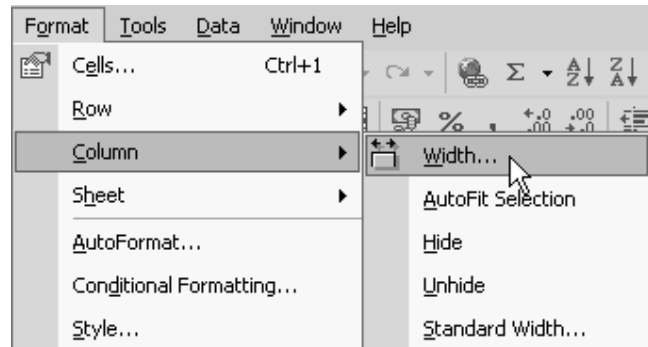
Hide / unhide columns

If you have a very wide spreadsheet, chances are that a number of columns contain “intermediate calculations” or less important information. Select the Hide option from the Format • Column menu to hide one (or more) selected column(s). The information is still there and the columns will retain their original names, but they are now hidden from normal view. You can *unhide* columns again by highlighting the column(s) to the left and right of the hidden columns and selecting the Unhide option from the Format • Column menu.

Changing width of several columns

Cell width is measured by the number of digits that would fit in a cell using the *standard* cell font (size).

You can set the width of several columns at once.



1. Select a number of columns by dragging the mouse cursor across the *column names* or select *all* columns by selecting the entire spreadsheet (click the top left square, to the left of the column A heading).
2. Set the width of the selected columns by using the Format • Column • Width... menu option.

Standard column width

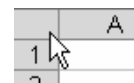
The *standard* or *default column width* is the width of the columns that have *not* been changed through one of the above operations. Set the *default width* of columns by selecting Format • Column • Standard Width...

Changing the height of rows

You change the height of a row, *mutatis mutandis*, in very much the same way you change the width of a column. Note that rows do not have a “default” height: they take the height of the largest *default* font (size) of any cell in the range (or larger if one of the row cells has been formatted to wrap around and the text consists of several lines).

Because of the wrapping of text, row heights can often end up varying. It can be useful to change all row heights to match your text height:

1. Select all rows by selecting the entire spreadsheet (click the select entire sheet square i.e. to the left of the column A heading).
2. Select Format • Row • AutoFit.



AutoFormatting a range

Adjusting the font size, style, alignment, borders, background and colours of the cells in your spreadsheet can be very time-consuming. Luckily, there is a smarter way of doing things: the **AutoFormat** option. AutoFormats are pre-defined templates (or models) that apply a whole lot of consistent formatting automatically to a selected range. Below are some sample *AutoFormats* applied to the *SA Sausage Sorcerer* example.

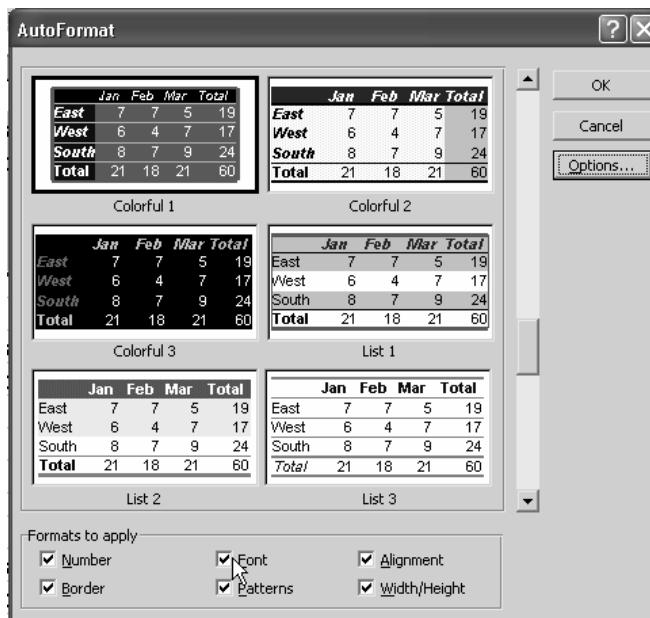
Classic Autoformat					3-Dimensional Autoformat				
	Jan	Feb	Mar	Sales		Jan	Feb	Mar	Sales
Boerewors (Standard)	R 3,400	R 3,210	R 3,385	R 9,995	Boerewors (Standard)	R 3,400	R 3,210	R 3,385	R 9,995
Morreesburg Special	R 2,145	R 2,250	R 2,640	R 7,035	Morreesburg Special	R 2,145	R 2,250	R 2,640	R 7,035
Pofadder Game Sausage	R 1,200	R 1,890	R 1,420	R 4,510	Pofadder Game Sausage	R 1,200	R 1,890	R 1,420	R 4,510
Willy's Wors	R 2,120	R 1,300	R 1,965	R 5,385	Willy's Wors	R 2,120	R 1,300	R 1,965	R 5,385
Sales:	R 8,865	R 8,650	R 9,410	R 26,925	Sales:	R 8,865	R 8,650	R 9,410	R 26,925

List Autoformat					Accounting Autoformat				
	Jan	Feb	Mar	Sales		Jan	Feb	Mar	Sales
Boerewors (Standard)	R 3,400	R 3,210	R 3,385	R 9,995	Boerewors (Standard)	R 3,400	R 3,210	R 3,385	R 9,995
Morreesburg Special	R 2,145	R 2,250	R 2,640	R 7,035	Morreesburg Special	R 2,145	R 2,250	R 2,640	R 7,035
Pofadder Game Sausage	R 1,200	R 1,890	R 1,420	R 4,510	Pofadder Game Sausage	R 1,200	R 1,890	R 1,420	R 4,510
Willy's Wors	R 2,120	R 1,300	R 1,965	R 5,385	Willy's Wors	R 2,120	R 1,300	R 1,965	R 5,385
Sales:	R 8,865	R 8,650	R 9,410	R 26,925	Sales:	R 8,865	R 8,650	R 9,410	R 26,925

Colourful Autoformat					List Autoformat				
	Jan	Feb	Mar	Sales		Jan	Feb	Mar	Sales
Boerewors (Standard)	R 3,400	R 3,210	R 3,385	R 9,995	Boerewors (Standard)	R 3,400	R 3,210	R 3,385	R 9,995
Morreesburg Special	R 2,145	R 2,250	R 2,640	R 7,035	Morreesburg Special	R 2,145	R 2,250	R 2,640	R 7,035
Pofadder Game Sausage	R 1,200	R 1,890	R 1,420	R 4,510	Pofadder Game Sausage	R 1,200	R 1,890	R 1,420	R 4,510
Willy's Wors	R 2,120	R 1,300	R 1,965	R 5,385	Willy's Wors	R 2,120	R 1,300	R 1,965	R 5,385
Sales:	R 8,865	R 8,650	R 9,410	R 26,925	Sales:	R 8,865	R 8,650	R 9,410	R 26,925

To apply an *AutoFormat*, follow these easy steps.

1. Highlight the range of cells that you wish to apply the format to. A range typically includes a top row and left column with header and descriptive information.
2. Select *Autoformat* from the *Format* menu.
3. Scroll through the list of available *Autoformats*. As you scroll, you will see an example of each format to enable you to visualise the final look.
4. Click the *Options* button if you only want to apply selected format elements e.g. only the alignment, patterns, colours or fonts.



Conditional Formatting

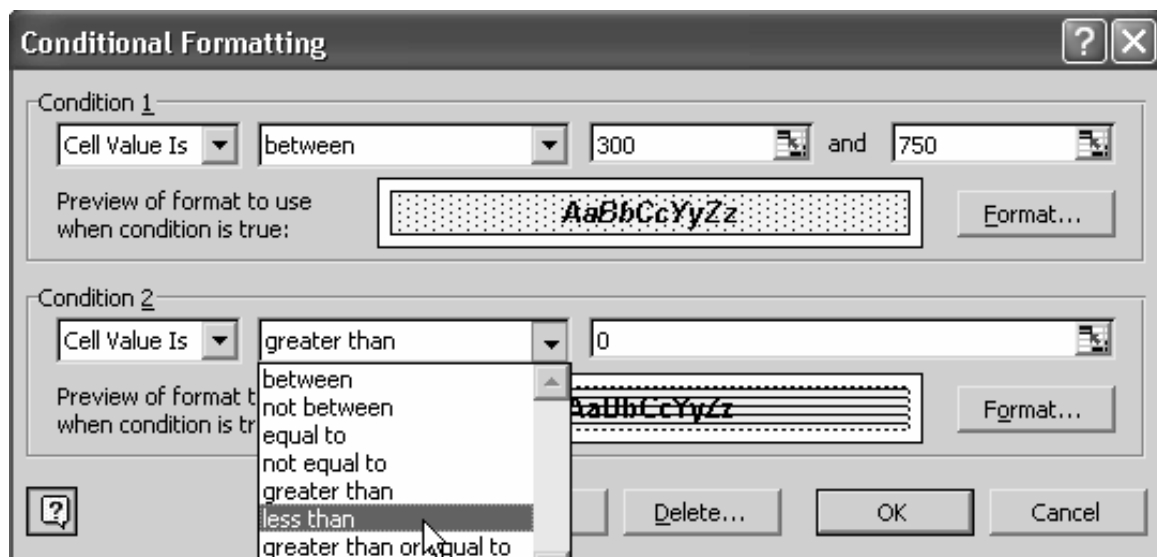
It is also possible to apply *different* formats to one single cell, depending on certain conditions. Usually the condition is related to the value in the cell. A typical example is where values outside an expected (or normal) range of acceptable values should be flagged very visibly e.g. by means of a red background. You can specify several conditions, each with their own special formatting rules.

To set a *Conditional Format*, perform the following steps.

1. Select (highlight) the range of cells to which you wish to apply the format.
2. Select the **Conditional Formatting...** option from the **Format** menu.
3. Specify the condition under which the conditional format must be applied by completing the required information boxes.
4. Click the **Format...** button to specify which format to use.

The example below applies a special format (bold, italic, yellow, thin border and dot pattern) to all cells within the selected range that contain values between 300 and 750.

Alternatively, the cells that contain a negative value (less than 0) will have horizontal lines on a light grey background with a dotted border.



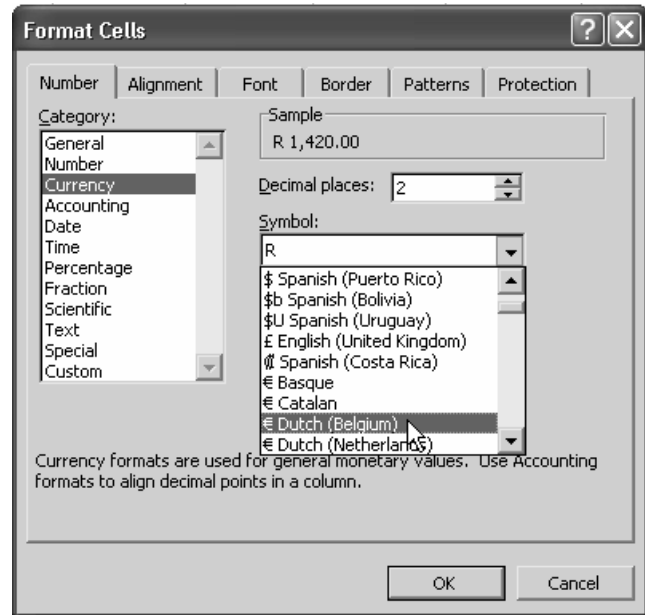
Number Formatting

Although number formatting is actually one of the cell formatting options, it deserves special consideration because of its importance. Most spreadsheets contain many numbers (and formulas) and it is important to realise that numbers can be formatted in different ways.

Common number formats

The following are the main number formats.

- **General:** displays as much of the number as possible.
- **Number:** displays a fixed number of decimal places. If a number has more significant digits after the decimal than need to be displayed, it gets rounded (for display only, not for calculation purposes). If it has less, zeros are added. E.g. both the numbers 123.44996 and 123.45 would be displayed as **123.4500** in number format with 4 decimal places.
- **Currency:** uses a currency symbol (R, \$, Fr, ...) and fixed number of decimal places (usually two or none). The currency symbol (and its placement before or after the amount) is determined by your system configuration, but the default value can be overridden for selected cells..
- **Accounting:** similar to the *Currency* format but with lots of space between the currency symbol and the number (decimals and currency symbols are vertically aligned).
- **Percentage:** displays a number as a percentage with a given number of decimal places. The number is multiplied by 100 and a %-symbol is added. E.g. 0.5 or $\frac{1}{2}$ equals 50% and 2.25 equals 225% in percentage format.
- **Fraction:** displays a number as a fraction. You can select one, two or three-digit fractions, or the standard (US) stock fractions of eighths and sixteenths. The integer part of the fraction will precede the fraction and is separated by spaces. For example 4.125 can be displayed as 4 $\frac{1}{8}$ and -2.33333333 as $-2 \frac{1}{3}$.
- **Scientific:** uses an exponential format to indicate the power of ten by which the number (mantissa) must be multiplied. This format is very useful for extremely large and very small numbers. 123,000,000,000 will be displayed as 1.23 E+11 (move the decimal point 11 places to the right or multiply with 1.23 with 10^{11}); 0.000000123 is equivalent to 1.23 E-07 (move the decimal point 7 places to the left or multiply with 10^{-7}).
- **Text:** treats a number as a text label instead of a value (e.g. 320 in BMW 320-series).



- **Custom and Special:** allow you to apply special formats to number-like codes e.g. telephone numbers, bank account numbers, product codes etc.

The following are illustrative examples for some of the above formats. The date and time formats will be discussed under a separate heading.

Number format	Example 1	Example 2
General format	3.141592654	-765.432
Number (4 decimals)	3.1416	-765.4320
Currency (2 decimals)	R 3.14	-R 765.43
Accounting (2 decimals)	R 3.14	R -765.43
Percentage (1 decimal)	314.2%	-76543.2%
Fraction (3 digits)	3 16/113	-765 54/125
Scientific (3 decimals)	3.142E+00	-7.654E+02
Custom (leading zeros)	00003	-00765
Special (Soc Sec No)	000-00-0003	-000-00-0765

What you see is *NOT* what you get

It is extremely important to note that the display format does *not* affect the internal accuracy or representation of the number. If a value of 1.2345 has a currency display format with 2 decimal places, it will appear as R1.23 on screen (and on print-outs) but for all calculations the *full* precision will be used. If another cell would multiply this value (say the price of an item as calculated by converting with the foreign exchange rate) by a factor two (representing the number of items ordered), the total value would be displayed as R2.47 (which is a rounded display for 2.4690), not R2.46 as might be expected (see example 2 below).

Excel typically uses a precision of 15 significant digits for all numbers in the spreadsheet, but numbers are usually only displayed with a limited precision. The spreadsheet extract below illustrates a few more examples. (Note: it *is* possible to set the *Calculation Options* to “Precision as displayed” i.e. “what you see is all the precision you get”.)

Example 1: adding fractions			Example 2: product prices			
fraction 1	1 1/3	1.33	Price of product in US\$:	0.25		
fraction 2	1 1/3	1.33	x US\$/Rand exchange rate	4.938		
fraction 3	1 1/3	1.33	= Price of product in Rand	R 1.23		
Total:	4	4.00	x Number of products sold	2		
			= Amount to be invoiced	R 2.47		
Example 3: student marks						
Student Name	Test 1	Test 2	Test 3	Test 4	Total	Pass/Fail
Joe Johnson	10.1	10.1	10.1	10.1	40%	Fail
Manual Mnabani	13.4	13.4	13.4	10.0	50%	Pass
Suzie Suzman	13.2	13.2	13.2	10.0	50%	Fail

Automatic Number Formatting

Excel will automatically apply certain number formats for you as you enter data. Here are a couple of examples.

Enter **3 3/5** and Excel will store the value 3.6 and apply the *Fraction, 1 digit* format.

Enter **3-Mar-98** and Excel will store it as a date and apply the date format (see later).

Enter **3:40** and Excel will convert it into a time and apply the time format.

Enter **R3.40** and Excel will apply the *Currency, 2 decimals* format.

Enter **R3,000** and Excel applies the *Currency, 0 decimals* format.

Enter **50%** and Excel applies the *Percentage (0 decimals)* format to the value 0.5.

Sometimes Excel may become a bit confused. If you enter **3/5**, Excel assumes that you want to enter a date (3rd March). If your intention was to enter the value three-fifths (0.6), you need to precede the fraction with a “dummy” integer value of 0 in order to “force” Excel into the fractional format i.e. you would enter **0 3/5** instead!


Excel Formulas

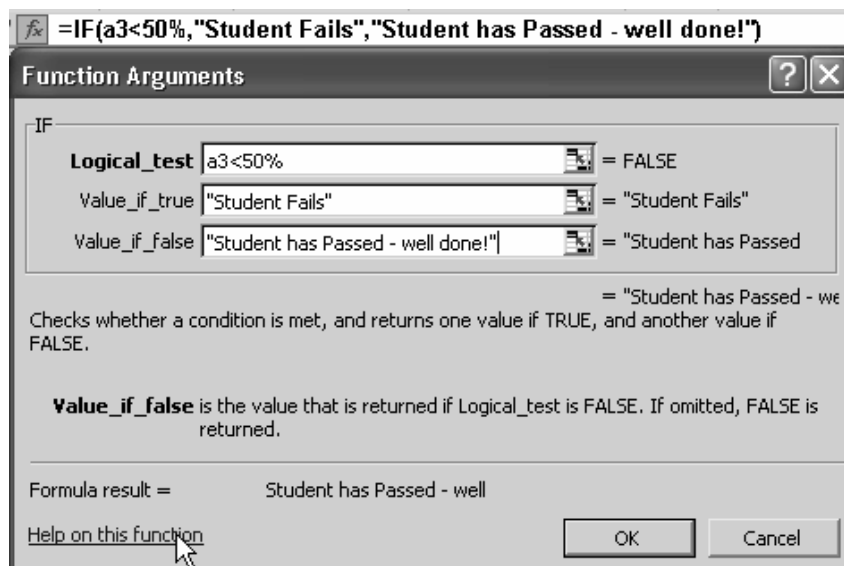
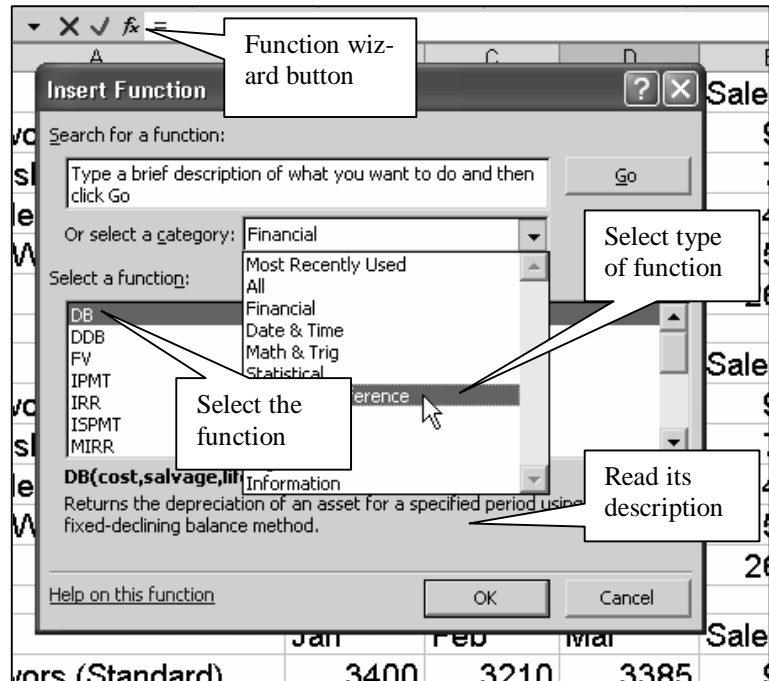
Excel's Built-in Functions: the Basics

Just like scientific and financial calculators offer much more advanced functionality than the mere basic arithmetic operations, Excel has a great number of **specialised functions** built in. The normal way of using the basic functions is by typing in their name and their arguments.

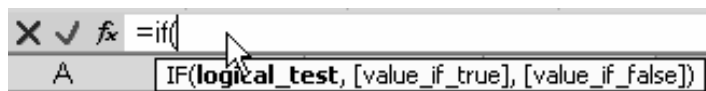
If you use a more complex or less frequently used function, you may not remember its exact name or argument.

Make use of the *function wizard* to assist you.

1. Click the *Insert Function* button 
2. Select the type of function from the *function category* menu
3. Click the *name* of the function you're interested in to see its description and the type of arguments it takes.
4. Excel will assist you in completing the function by prompting you for each of the arguments the function requires. Remember that you can move the *Function Arguments* window around if the cells you want to see are obscured.



If you do remember the name of the function, Excel will prompt you for the required and optional arguments as you type.



A function consists of the following elements:

- **Function name:** you can enter a name in lower or upper case but Excel will automatically convert it to upper case.
- **Round brackets:** this identifies it as a function.
- **Argument(s):** these are the values or cell references that the function uses to calculate its result. (Mathematicians refer to the arguments as the *independent variables*.) Most functions require more than one argument; if this is the case, the arguments usually have to be in a strict order. Arguments are separated by commas.

An example of a function is:

`=ROUND(A5,2)`

This function will round (hence the name of the function) the value in cell A5 (the first argument) to the nearest one-hundredth (two “decimal places” to the right of the decimal point, as indicated by the second argument) and place its result (the rounded value) in the current cell.

Note that the arguments of functions can themselves be any formula or function. Using a function within a function is called *nesting*, as in:

`=B7+AVERAGE(MIN(A2:A5),MAX(B2:B5))`

This formula first finds the lowest value in range A2:A5, then calculates the highest value in B2:B5 to B7, averages the two and adds the result to B7.

Some formulas can be nested extremely deeply, which makes them almost impossible to review or change at a later date. If you find that you have one of these very complex formulas, consider breaking up the formula into smaller components by storing some intermediate calculations in other cells.

The best way to get an overview of the variety and range of functions that are available in Excel is by browsing *Help* or by using the *function wizard*. We will briefly introduce some of the functions that you are likely to use in a business context.

Statistical functions

Statistical functions are used to summarise large amounts of numbers. The following are the more basic statistical functions. The sample values are based on the data set that is highlighted.

	A	B	C	D	E
1	Dataset (A2:C4)				
2	3	4	10		
3	4		0		
4	7	1	5		
5					
6	Function	Value for (A2:C4)	Description	Notes	
7	SUM()	26	Sum total of all values	Use the AutoSum toolbar button	
8	AVERAGE()	3.25	Average (=sum/count)	Includes zero values but not empty cells	
9	COUNT()	8	Number of values	Doesn't count empty cells	
10	MAX()	10	Highest value in range	See also the "large" function	
11	MIN()	-4	Lowest value in range		
12	VAR()	18.786	Variance of all values		
13	STDV()	4.334	Standard deviation		
14	MEDIAN()	3.5	Middle value of the set	For an even number of values: the average of middle two values (here: 3 and 4)	

Mathematical functions

Although you are unlikely to use any of the advanced mathematical functions such as those dealing with imaginary numbers or trigonometry, business spreadsheets often need a way of dealing with decimal fractions and rounding values. The following are some of the functions that you will sooner or later deal with and sample values (based on A1 = 1234.563).

	A	B	C	D	E
1	1234.563				
2					
3		ROUND(x,y)	INT(x)	ABS(x)	CEILING(x,y)
4	Description:	Rounds x to y decimals.	Returns the integer < or = to x	Converts x to a positive number	Rounds x to a multiple of y (away from 0)
5	Example 1	=ROUND(A1,0)	=INT(A1)	=ABS(A1)	=CEILING(A1,1)
6	Value:	1235	1234	1234.563	1235
7	Example 2	=ROUND(A1,2)	=INT(-A1)	=ABS(-A1)	=CEILING(A1,0.05)
8	Value:	1234.56	-1235	1234.563	1234.6
9	Example 3	=ROUND(A1,-2)	=A1-INT(A1)		=CEILING(A1,200)
10	Value:	1200	0.563		1400
11					
12	Notes:	See also MROUND()		See also SIGN()	See also FLOOR()

You should remember that *What-You-See-Is-NOT-Always-What-You-Get*. A number may be displayed with only 2 decimals but the real underlying value that is used for calculations may have far more precision. Refer to the discussion earlier in this document. This is why you may need some of these functions – especially the *ROUND*-function is handy!

Another very useful function is the **random number generator**. You can use this function when testing a spreadsheet model structure without having access to actual data (yet) or to generate forecast data to test out various business scenarios. You generate random numbers by means of the following formulas:

=RAND() : generates a random number between 0 and 1

=RANDBETWEEN(lower limit, upper limit) : generates an integer random number between the specified lower and upper limits. Note that the RANDBETWEEN function is only available if you have the *Analysis Toolpak Add-in* installed.

Random numbers are re-generated each time a(nother) cell value in the spreadsheet changes. You can also “force” new random numbers to be generated by pressing the *Recalculation* key: the **<F9>** function key.

	A	B
15	Jean-Paul's Excel-nt Casino	
16	Your throw:	
17	Dice 1:	2
18	Dice 2:	2
19	To throw again, press <F9>	

On the left is a more playful example.

Each time **<F9>** is pressed, two new numbers (between 1 and 6) appear in cells B17 and B18. Here I was lucky to throw double-two! You can make dice that throw any range of numbers.

Notes:

The *RAND* function has brackets but doesn't require any arguments between the brackets.

Don't confuse the *RAND* function name with the South African currency – the similarity is just a coincidence.

Where the *RANDBETWEEN* function is not available, you can use the following equivalents:

- **=RAND()*range+base** : will generate a number between base and (base+range)
- **=INT(RAND()*range+base+1)** : will generate *integer* numbers.

E.g. to generate the dice numbers above the function **=INT(RAND()*6+1)** was used:

- **RAND()** generates (evenly distributed) numbers between 0 and 1
- **RAND()*6** generates numbers between 0 and 6
- **RAND()*6+1** generates numbers *between* 1 and 7 (but never quite 1 or 7!)
- **INT(RAND()*6+1)** returns the integer part of the above number i.e. numbers between 1 and 6. This is the function used in the casino example above.

The Excel IF function

Arguably the most often used logical function is the **IF** function. This function allows you to return any one of two values, depending on the outcome of a condition.

The **IF** function has three arguments:

- The condition (or criterion) that has to be checked. You make use of the following operators to construct your condition: = (equal), > (greater than), < (less than), <> (not equal), >= (greater than or equal), <= (less than or equal).
- The value that will be returned if the condition holds true; the value can be any formula.
- The value that will be returned if the condition does *not* hold true i.e. is false.

The following are some examples of simple applications of the **IF** function.

Example 1: calculating tax on positive income only

A company will only be liable for income tax if it earns a positive income. Assuming that the tax rate is 40% and the income is contained in cell D20, the following formula would calculate the income tax due (spaces added for legibility):

=IF (D20 > 0, D20 * 40%, 0)

This should be read: If (income exceeds 0) then (tax = income * 40%) else (tax = 0).

An alternative but equally valid way would be:

= IF (D20 < 0 , 0, 40% * D20)

If (income is negative) then (tax = nil) else (tax = 40% of income)

Example 2: calculating a pass or fail

Assuming that 50% is required for a student to pass, the following formula will calculate the Pass/Fail symbol (assuming the student mark is in E10):

=IF (E10 >= 50% , "Pass" , "Fail")

If (mark is greater than or equal to 50%) then (symbol = Pass) else (symbol = Fail).

Note that, since the value of the symbol is text rather than a number, the symbols are enclosed in double quotes.

If you want *nothing* to be returned if a student passes, you could use the formula:

=IF (E10 < 50%, "Fail")

This would display nothing for students who obtain at least 50% and flag only the *Fail students*.

More complex conditions

More complex conditions can be expressed by using the logical functions AND() and OR() or by using *nested conditions*. The following are some examples:

=IF (OR (E2 < 50%, F2 < 50%, G2 < 50 %), "Fail", "Pass")

If (E2 < 50% or F2 < 50% or G2 < 50%) , then fail, else pass. You fail if you have less than 50% in either E2, F2 or G2 (or more than one). You only pass if you have 50% or more for all three. This is quite different from:

=IF (AND(E2 < 50%, F2 < 50%, G2 < 50 %), "Fail", "Pass")

Here you pass as soon as you have at least 50% for *any one* of E2, F2 or G2 i.e. only one single 50% score is sufficient to secure a *Pass*.

The nested *IF* allows you to have more than two options.

=IF (D2 < 50%, "Bad", IF(D2 >=70%, "Excellent", "Good"))

If D2 is less than 50%, then you get a "Bad" symbol, If D2 is at least 70%, you get an "Excellent" symbol, else (if D2 is at least 50% but less than 70%) you get a "Good".

When you get too many nested *IF* conditions, you should consider using the *LOOKUP* function. Compare the two following statements that aim to assign a symbol from A to F based on the final mark stored in D2:

=IF (D2 < 40%, "F", IF (D2 < 50%, "E", IF (D2<60%, "D", IF (D2 < 70%, "C" , IF (D2 < 75%, "B", "A")))))

(Note: the order is very important!). A much neater and less error-prone formula is:

=VLOOKUP (D2, Symbols, 2)

where *Symbols* is the range that holds a list of symbols and the corresponding mark cut-off.

Lookup and reference functions

A common business situation is where values have to be looked up in tables. Excel offers a number of functions that will enable you to choose a particular value from a whole list or set of values.

- **=CHOOSE(choice,option1,option2, ...):** allows you to choose a value from a list (option1 etc.) depending on the value of your choice argument (0, 1 ...)
- **=VLOOKUP(lookup value, range reference, column number):** locates a value (the lookup value) in the first column of a columnar table (range reference) and selects the corresponding value in the nth column (column number).
- **=HLOOKUP(lookup value, range reference, row number):** locates a value (the lookup value) in the first row of a horizontally organised table (range reference) and selects the corresponding value in the nth row (row number).

These functions are best illustrated by means of examples.

Example 1: UCT Parking Fees.

On the left, you find a simple table with the parking fees for a particular organisation. The

	A	B	C
2	Example 1:UCT Parking Fees		
3			
4	Code	Category	Parking Fee
5	1	Students	R 180.00
6	2	Staff	R 250.00
7	3	Academics	R 320.00
8	4	VIP	R 580.00
9	5	Disabled	R 200.00
10			
11	Enter code:	4	
12	Category:	VIP	
13	Parking Fee:	R 580.00	

fees vary according to the status of the individual, which can be referenced by means of a code. E.g. a student has a code of 1 and is liable for an annual parking fee of R180, whereas academics (code 3) pay R320 annually.

An Excel function can be constructed to quickly “look up” the parking fee amount due, based on the code entered by the user. Since the codes used here are sequentially numbered, there are two basic ways of constructing the formula in B13.

Using the CHOOSE function:

(B13) =CHOOSE(B11,180,250,320,580,200) or, slightly better:

(B13) =CHOOSE(B11,C5,C6,C7,C8,C9)

Depending on the value in B11, the CHOOSE-function selects (chooses) one of the amounts (values) which follow. The first version of the formula is not very desirable: when UCT’s management decides to increase the parking fees, you would not only need to update the schedule but also change the formula(s)! A problem with the CHOOSE function in general is that it requires “sequential” codes starting from 1. Another problem is that the formula can become quite long if there are many possible options. There is a better alternative.

Using the VLOOKUP function:

(B13) =VLOOKUP(B11,A5:C9,3)

This formula should be read in the following way: look the value in B11 up in the first column of the table A5:C9 and return the corresponding value (i.e. on the same row) in the 3rd column.

Since this example is perhaps a bit too trivial, let’s illustrate the lookup-function with a more realistic example.

Example 2: The South African income tax rates (for natural persons)

The table on the next page lists the income tax rates for natural persons in South Africa for the 1997 tax year. (The rates are adjusted annually but the principle remains.) South Africa has a system of progressive taxation: as your taxable income increases, you pay a higher tax rate.

For example, if your taxable income is R34,500, you would pay tax as follows:

- 17% tax on the first R15000 (or R2,550),
- 19% on the next R5,000 (tax=R950),
- 21% on the next R10,000 (tax=R2,100), and finally
- 30% on the last (highest) R4,500.

	D	E	F	G	H	I	J
2	Example 2: SA Income Tax Rates						
3	Taxable Income		Rates of tax:				
4	exceeds:	but less than:	Fixed Tax Portion	Margi nal Rate	on the amount in excess of:		
5	R 0	R 15,000	R 0	17%	R 0		2) In this tax table
6	R 15,000	R 20,000	R 2,550	19%	R 15,000		
7	R 20,000	R 30,000	R 3,500	21%	R 20,000		
8	R 30,000	R 40,000	R 5,600	30%	R 30,000		3) in this row
9	R 40,000	R 60,000	R 8,600	41%	R 40,000		
10	R 60,000	R 80,000	R 16,800	43%	R 60,000		
11	R 80,000	R 100,000	R 25,400	44%	R 80,000		
12	R 100,000		R 34,200	45%			
13							
14	Enter Taxable Income:		R 34,500	1) Lookup this income			
15	Fixed tax portion:		R 5,600	=VLOOKUP(G14,E4:I12,3)			
16	Marginal rate:		30%	=VLOOKUP(G14,E4:I12,4)			
17	on excess:		R 4,500	=G14-VLOOKUP(G14,E4:I12,5)			
18	Marginal amount:		R 1,350	=G16*G17			
19	Total tax (fixed+margl):		R 6,950	=G15+G18			
20							

Since the first three steps will apply to *any* taxable income between R30,000 and R40,000 these amounts have already been “pre-calculated” (R2,550+R950+R2,100=R5,600) and this can be read off the table. You just need to calculate the marginal tax for the highest tax *bracket*, for the amount in excess of R30,000.

The formulas in G15 to G19 calculate the tax in two separate steps.

- **G15** looks up the taxable income (as given in G14) in the first column of the entire tax table (E4:I12). Excel’s *LOOKUP*-function will go down the values in the first column (which should be in increasing order!) until it finds an exact match, or stops before the next larger value. In our case, it will “stop” in row 8 since the value in row 9 (E9) is R40,000 i.e. exceeds the lookup value. It will now move across the same row to the *third* column (the last argument of the function = 3) and return this value, which is R5,600, or the “fixed” portion of tax due for any taxable income between R30,000 and R40,000.

- **G16** looks up the *marginal tax rate* for the portion of the taxable income exceeding the R30,000 (the lower break of this tax bracket) in the same manner (except that it is located in the 4th column).
- **G17** calculates the amount on which the marginal rate should be calculated i.e. by *how much* the taxable income exceeds the lower break of the applicable tax bracket (or “band”).
- **G18** calculates the actual marginal tax due i.e. G16 x G17
- **G19** calculates the *total tax* due, i.e. the sum of the tax in G15 and G18.

Below are two more examples – see if you can work out how the *VLOOKUP*-function arrived at its values.

	D	E	F	G
13				
14		Enter Taxable Income:		R 62,000
15		Fixed tax portion:		R 16,800
16		Marginal rate:		43%
17		on excess:		R 2,000
18		Marginal amount:		R 860
19		Total tax (fixed+margl):		R 17,660

	D	E	F	G
13				
14		Enter Taxable Income:		R 150,000
15		Fixed tax portion:		R 34,200
16		Marginal rate:		45%
17		on excess:		R 150,000
18		Marginal amount:		R 67,500
19		Total tax (fixed+margl):		R 101,700

Notes:

- The *HLOOKUP* function works in a similar way, except that it assumes that the table is organised in rows rather than columns. The lookup value will be checked along the first row of the table (left to right) and the function will return the corresponding entry in the nth row of the same column.
- Some other spreadsheet programs start numbering columns (or rows) at 0 instead of 1.
- The first column (or row) of the table should be in increasing (or *ascending* order) – you may wish to use Excel’s *data sort* feature if your data isn’t organised sequentially. (Use the *Sort Ascending* button on the standard toolbar.)
- The *LOOKUP* function looks for values that equal or exceed the lookup value (but are less than the next value). Sometimes you require an *exact* match with the value in the table (e.g. when looking up product codes, names or student registration numbers): the *LOOKUP* function allows you to specify a(n optional) *fourth* argument *FALSE* if you require the exact match.

`=VLOOKUP("BNDJAM007",A2:F850,4,FALSE)`

will only return a value if there is a student with the registration number BNDJAM007 in the A column.

Convert Calculated Values in Constants Using Paste Special

Sometimes, it is necessary to turn the results of a formula into *permanently fixed constants* (e.g. when formulas refer to values in other spreadsheet files which may not be available later or if you want to reduce calculation time/memory requirements). First *Copy* the range with the formulas whose values you wish to make permanent (with **Edit • Copy** or **<Ctrl>-<C>**) and then select the **Paste Special** option from the **Edit** menu. You can now elect to copy only the values (or the formatting) from the clipboard to another area of the spreadsheet or over the original cell formulas. If you have used very complex formulas, you are well advised to keep a copy of the formula somewhere in the spreadsheet (e.g. as a text entry).

More on Using Logical Conditions in Formulas

Logical conditions were discussed with the IF statement. It is possible to use the way Excel evaluates logical conditions in different ways.

The following example shows how Excel assigns a *truth value* to a logical condition.

	A	B	C	D	E	F	G
1			This formula:	results in:		This formula:	results in:
2	3		= (A2 > A3)	FALSE		= D2 + 1	1
3	5		= A3 > A2	TRUE		= D3 + 1	2
4			= A3 = A2	FALSE		= D4 + 1	1
5			=A2=A3=A4	TRUE		= D5 + 2	3
6							

With A2 assigned the value 3 and A3 the value 5, the formulas in column D (shown in formula form in column C) evaluate to either TRUE or FALSE.

Note in particular the funny formula “=A3=A2 “. In fact, the following is a perfectly legal formula: =(A2<A3)=(A3<A2) which evaluates as FALSE since the truth value of the statement A2<A3 is not the same as the truth of the statement A3<A2.

Internally, Excel assigns a numeric value of **1** to a TRUE condition and a numeric value of **0** to a FALSE condition.

Hence the somewhat perplexing truth of the last formula “=A2=A3=A4”. If there are no brackets, Excel evaluates from left to right, thus A2=A3 evaluates as FALSE and internally stored as **0**. Since A4 is an empty cell, it is considered to be **0** in formula calculations and thus 0 = 0 which is true!

You can therefore use evaluated conditions in formulas as shown in column G whereby a number is added to the truth values in column D. This feature enables you to write some IF formulas differently e.g. the tax formula given earlier to calculate tax on positive income = IF (D20 > 0 , 40% * D20 , 0) could be rewritten as =(D20>0)*D20*40%. Note that formulas like these are very difficult to understand or check, so we advise that you should avoid using them as much as possible.

Mixed addressing

In the introductory module, the difference between absolute and relative cell referencing was introduced. There exists a hybrid form between absolute and relative cell addresses. This is called a *mixed* reference and, though it takes quite a bit of effort to understand, it may save you invaluable time when building more complex spreadsheets.

The following example calculates how much R1000 will be worth when it is invested for a various number of years (read off vertically) and at different interest rates (read off horizontally). It is assumed that you leave the interest (which we'll calculate on an annual basis for simplicity's sake) in your account: this is called *compounding*. The power behind compounded interest is that, in the following year(s), you *earn interest on previously accumulated interest*. While the initial additional earnings are indeed quite small, they tend to build up to dramatic amounts after ten or more years. The formula used for compound interest is the following:

$$\text{Balance after } n \text{ years} = \text{Initial capital} \times (1 + \text{interest rate})^{\text{number of years}}$$

Although this formula looks frightening, it is actually not all that different from the formula that is used for simple interest i.e. where no interest is earned on previous interest.

$$\text{Balance after } n \text{ years} = \text{Initial capital} \times (1 + \text{interest rate} \times \text{number of years})$$

		= \$G\$1*(1+B\$2)^\$A3					
	A	B	C	D	E	F	G
1	Compound Interest Table			Initial Capital:			R 1,000
2	No Years	8%	10%	12%	14%	16%	18%
3	1	R 1,080.00	R 1,100.00	R 1,120.00	R 1,140.00	R 1,160.00	R 1,180.00
4	2	R 1,166.40	R 1,210.00	R 1,254.40	R 1,299.60	R 1,345.60	R 1,392.40
5	3	R 1,259.71	R 1,331.00	R 1,404.93	R 1,481.54	R 1,560.90	R 1,643.03
6	4	R 1,360.49	R 1,464.10	R 1,573.52	R 1,688.96	R 1,810.64	R 1,938.78
7	5	R 1,469.33	R 1,610.51	R 1,762.34	R 1,925.41	R 2,100.34	R 2,287.76
8	6	R 1,586.87	R 1,771.56	R 1,973.82	R 2,194.97	R 2,436.40	R 2,699.55
9	7	R 1,713.82	R 1,948.72	R 2,210.68	R 2,502.27	R 2,826.22	R 3,185.47
10	8	R 1,850.93	R 2,143.59	R 2,475.96	R 2,852.59	R 3,278.41	R 3,758.86
11	9	R 1,999.00	R 2,357.95	R 2,773.08	R 3,251.95	R 3,802.96	R 4,435.45
12	10	R 2,158.92	R 2,593.74	R 3,105.85	R 3,707.22	R 4,411.44	R 5,233.84
13	11	R 2,331.64	R 2,853.12	R 3,478.55	R 4,226.23	R 5,117.26	R 6,175.93
14	12	R 2,518.17	R 3,138.43	R 3,895.98	R 4,817.90	R 5,936.03	R 7,287.59
15	13	R 2,719.62	R 3,452.27	R 4,363.49	R 5,492.41	R 6,885.79	R 8,599.36
16	14	R 2,937.19	R 3,797.50	R 4,887.11	R 6,261.35	R 7,987.52	R 10,147.24
17	15	R 3,172.17	R 4,177.25	R 5,473.57	R 7,137.94	R 9,265.52	R 11,973.75

The compound interest table can be built very quickly in Excel. (Use *AutoFill* to complete row 2 and column A; and *AutoFormat* if you want a more attractive formatting of the table).

The element of most interest (excuse the pun) in the table is the formula that calculates your final balance given the interest (in row 2) and the number of years you invested (in column A). This formula was only entered once and then copied across the entire table range. Below is an extract of a portion of the same spreadsheet in *formula* view.

	A	B	C	D
1	Compound Interest			
2	No Years	0.08	0.1	0.12
3	1	=G\$1*(1+B\$2)^\$A3	=G\$1*(1+C\$2)^\$A3	=G\$1*(1+D\$2)^\$A3
4	2	=G\$1*(1+B\$2)^\$A4	=G\$1*(1+C\$2)^\$A4	=G\$1*(1+D\$2)^\$A4
5	3	=G\$1*(1+B\$2)^\$A5	=G\$1*(1+C\$2)^\$A5	=G\$1*(1+D\$2)^\$A5
6	4	=G\$1*(1+B\$2)^\$A6	=G\$1*(1+C\$2)^\$A6	=G\$1*(1+D\$2)^\$A6
7	5	=G\$1*(1+B\$2)^\$A7	=G\$1*(1+C\$2)^\$A7	=G\$1*(1+D\$2)^\$A7

We will analyse the formula in B3 step by step.

- The base formula is $\text{= Initial Capital} * (1 + \text{interest rate}) ^ \text{Number of Years}$
- Using relative referencing only, this would be $\text{= G1} * (1 + \text{B2}) ^ \text{A3}$ for B3.
- We have entered it using mixed referencing as $\text{= } \$\text{G}\$1 * (1 + \text{B}\$2) ^ \text{\$A3}$ for B3.

Initial Capital is in cell G1. Since this reference should **always** be to G1, no matter where the formula is copied to, it should be an absolute reference i.e. **\$G\$1**.

The interest rates can *always* be found in *row 2*, irrespective of the row to which the formula will be copied, hence the reference to row 2 should be absolute. However, for the column C, the appropriate interest rate is *not* **B2** (8%) but **C2** (10%); and the interest rate for column D is 12% i.e. **D2**. Hence, when the formula is copied across different columns, the *column* reference should change accordingly, and therefore the *column* reference must be relative. This is precisely a case for *mixed* cell referencing: **B\$2**. Check for yourself how the B\$2 changes to C\$2, D\$2 etc when moving (copying) from column to column but never changes when going down the rows: the **\$2** is the absolute (or unchangeable) part of the formula.

Exactly the opposite is true for the reference to the number of years. These can be found in column A. For B3, the number of years is 1 i.e. A3. The reference to column A should never change when copied across from column to column, hence this part of the address should be fixed or made absolute: **\$A**. However, for row 4, the number of years is 2 and the reference should change to **A4**, the row reference must be relative i.e. *no* \$-sign. Check for yourself how **\$A3** changes to \$A4, \$A5 etc when moving (copying) from row to row but never changes when moving across columns: the **\$A** is the absolute part of the mixed cell reference.

Using Headers as Cell Range Names

There is still another, much more intuitive but sometimes quite confusing way of referencing cells. You will remember that, instead of having to type in cell addresses when entering a formula, you can avoid typing mistakes by just clicking the cell you wish to insert; this is called the *pointing* method. Nevertheless, when reading formulas it is still not always imme-

diately clear what the cell references mean. There should be an easier way to read the formulas, and there is! Refer to the simple spreadsheet which was introduced in the introductory module to illustrate the concepts of relative and absolute addressing.

	A	B	C	D	E	F
1	Jean-Paul's Excel-nt Supermarket					
2				VATRate:	14%	
4	Item	Quantity	Price	Total (excl)	VAT	Total (incl)
5	Apples	3.5	R 3.49	=Quantity*Pr		R 12.22
6	Bananas	1.25	R 2.45	R 3.06	R 0.43	R 3.49
7	Cherries	0.75	R 7.45	R 5.59	R 0.78	R 6.37
8			Total:	R 20.87	R 1.21	R 22.08

Instead of entering the formula (D5): =B5*C5, you can enter the much more legible and intuitive formula =quantity*price. Is Excel clever enough to figure out what you want? YES! Similarly, you can enter the following:

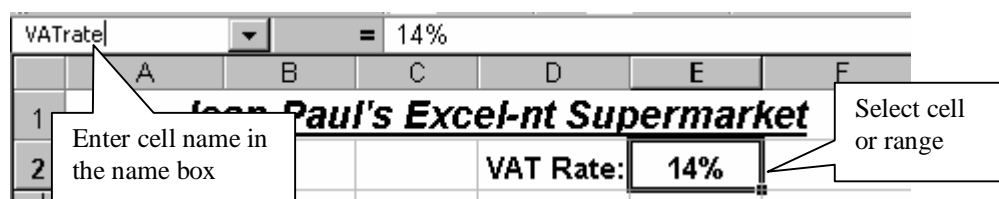
(E5) =Total(excl)*\$E\$2 (You still have to refer to the VAT Rate as \$E\$2)

(F5) =Total(excl)+VAT

You can even enter the following in E8: instead of =SUM(E5:E7)

(E8) =SUM(VAT)


Excel is normally sufficiently clever to infer the cell reference from a descriptive cell name by using the context i.e. from column and row labels. Sometimes, however, you may prefer to give Excel explicit instructions about names: you can assign a **name** to any cell or range of the spreadsheet. In fact, it is advisable to always declare your cell names using a formal procedure to avoid duplicate names and/or ambiguities. The quickest way to assign a name to a cell or range is highlight the cell (or cell range), click on the **name box** on the left of the formula bar and type in your desired name.



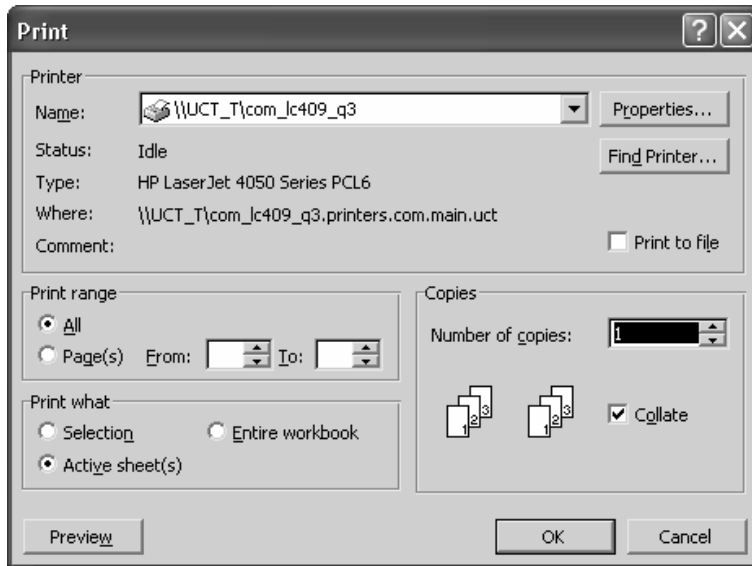
For the *invoice* example, we assign the name *VATRate* (capitalisation is not important) to the cell E2, so we could actually use the formula =Total (excl)*VATRate in E4:E6.

And there is an important added bonus: *we don't have to worry about absolute addressing in the formula any more!* Cell and range names make your spreadsheet formulas much easier to read. You can also use the names whenever a command requires a range from you, e.g. for graphs or printing.

Print Options

To print your spreadsheet using the default settings, click the *Print* button  to print out the current spreadsheet on your default printer. Note that hovering your mouse cursor above the button will display your default printer.

The Print dialogue window

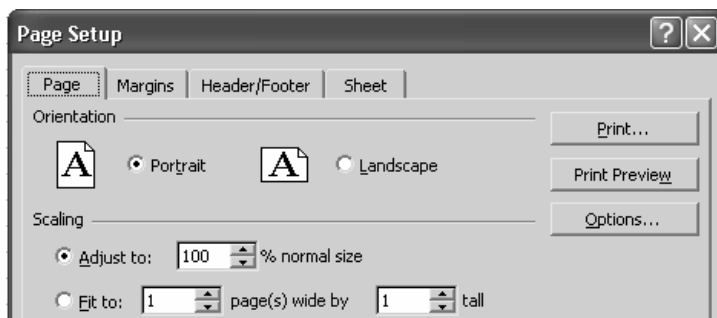



To change print defaults, press <Ctrl>-<P> or select the **File • Print...** menu option to bring up the standard *Print* dialogue window. This will allow you to change the number of copies, the printer and which pages to print.

You can also select a particular range or a number of workbook sheets (<Ctrl>-click the required sheet tabs) prior to bringing up the *Print* window.

Changing the Page Layout

You will often want to customize the layout of your spreadsheet printout. To access the *Page Setup* dialogue window from either:



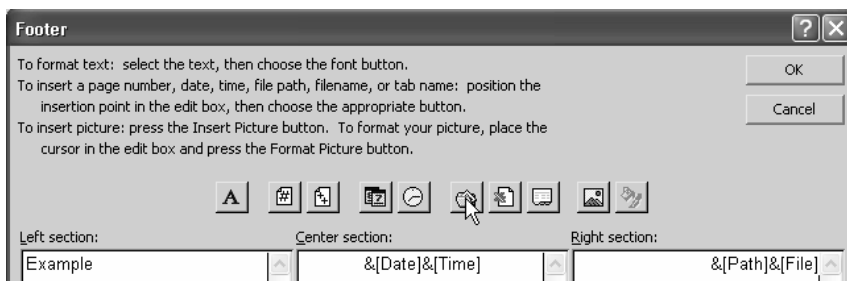
- Click the *Print Preview* button  and click *Setup...* from the preview screen, or
- Select the **File • Page Setup...** menu option.

If you are working with a large spreadsheet, you will often want to

scale the print-out, either to a fixed percentage of normal size, or else fit the print-out to a specified number of pages. Many spreadsheets are too wide to fit on a *portrait*-orientated page, so you will often set the default paper orientation to *Landscape*.

Adding Headers or Footers

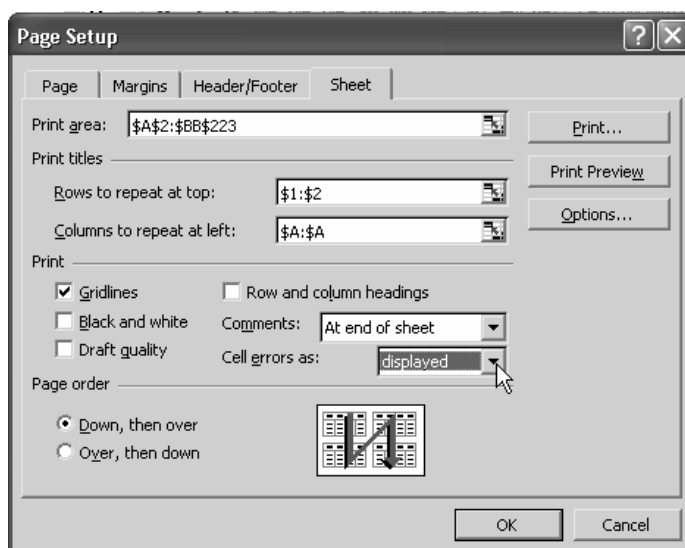
Click the *Header/Footer* tab to print a common header or footer on each page. Apart from including standard text, you can customize the header/footer to include page numbers, file names, date and/or time printed etc.



Setting Sheet Options

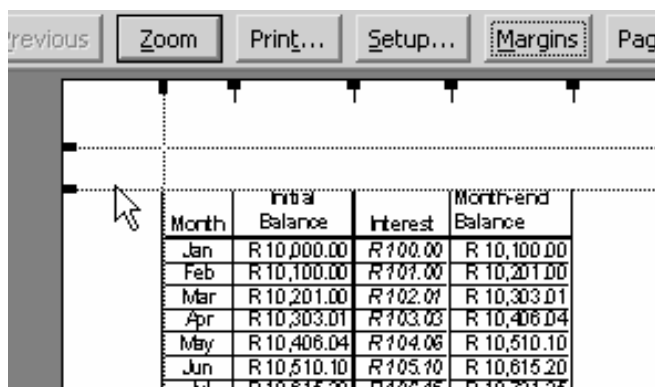
Extremely useful are the different *Sheet* options of the *Page Setup* window that allow you to:

- *repeat* certain columns and/or rows on each page
- print the faint *grid lines* between the cells
- change the cell range (*print area*) to be printed.



Adjusting Print Margins

The print margins can also be defined from the *Page Setup* window or, more visually, by clicking the *Margins* button in the *Print Preview* window and dragging the margin borders as desired.



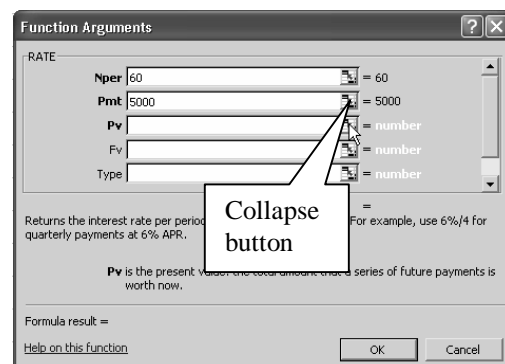
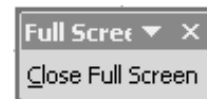
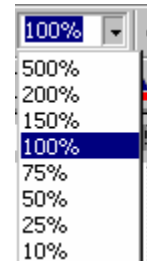
The *Print Preview* window also allows you to adjust the column widths for printing purposes only, without changing the actual column widths of your spreadsheet. Drag the column border handles at the top of the page to change the column widths.

Viewing large spreadsheets

As your spreadsheets become larger and larger, you may not be able to see all of the relevant cells in the same window. There are a number of different ways in which you can remedy this situation.

Most of the following methods to see more of your document, apply to other applications (such as *MS-Word*), too.

- Change the window zoom factor on the *standard toolbar* (see right). Select any factor from the drop-down box or enter your own % factor. (This option can also be selected from the View menu.) The smaller the %, the more you can see, e.g. 50% shows much more than 100% but the text may be too small to be readable.
- If your document or application window is not maximised, click the *maximise button*.
- You can also (temporarily) remove all scroll bars, toolbars, title bar and status bar by selecting Full Screen from the View menu. Only the *menu bar* will remain on-screen and the rest of your “screen estate” will be devoted to your spreadsheet. To switch back to your normal *editing screen*, click the Close Full Screen option on the small “floating” *Full Screen* window (you can move this window out of your way by dragging its title bar).
- Some *pop-up* windows (especially those generated by the wizards) require you to enter data but may be obscuring your view. It is not always feasible to move them out of your way by dragging the title bar due to their size. These windows often have a *collapse* button that removes the window except the data entry (dialogue) box.



Splitting document windows: you can split the document window both horizontally and vertically to view different (rectangular) areas of your *same* document: these areas are called **panes**. Do this by dragging the little bar at the top (right) end of the vertical (horizontal) scroll bar down

	A	B	C	I	J	K
2	No Years	8%	10%	22%	24%	26%
3	1	R 1,080.00	R 1,100.00	R 1,220.00	R 1,240.00	R 1,260.00
4	2	R 1,166.40	R 1,210.00	R 1,488.40	R 1,537.60	R 1,586.80
5	3	R 1,259.71	R 1,331.00	R 1,815.85	R 1,906.62	R 1,997.39
6	4	R 1,360.49	R 1,464.10	R 2,215.33	R 2,364.21	R 2,513.09
23	21	R 5,033.83	R 7,400.25	R 65,096.32	R 91,591.55	R 128,086.78
24	22	R 5,436.54	R 8,140.27	R 79,417.51	R 113,573.52	R 161,555.01
25	23	R 5,871.46	R 8,954.30	R 96,889.36	R 140,831.16	R 203,023.24
26	24	R 6,341.18	R 9,849.73	R 118,205.02	R 174,630.64	R 256,491.47
27	25	R 6,848.48	R 10,834.71	R 144,210.13	R 216,541.99	R 323,001.70
28	26	R 7,396.25	R 11,919.40	R 175,026.26	R 269,543.07	R 407,011.91

(left) to the desired position. (This is similar to *Word's* split windows.) You can also select the Split option from the Window menu. Double-clicking any of the window splits will remove the split window. Note

how this split window allows you to check how much your R1,000 will be worth after 25 years, invested at 24% p.a.

Working with dates and time

Many spreadsheets incorporate dates and, to a lesser extent, time. Often you will want to do calculations with dates such as adding 7 days to the current date to calculate an expected delivery date. Excel has a very interesting way of dealing with date and time arithmetic.

The underlying philosophy for date manipulation rests on the following principles:

- Excel counts days as consecutive numbers. Each day (date) is represented by a unique number.
- Excel started counting days from the 1st Jan, 1900, which can be considered “Day 1”. 2nd Jan, 1900 was day 2; 31st Dec 1900 was day 366; 1st Jan 1901 was day 367; 1st April 1998 is day 35886; 2nd April 1998 is day 35887; 1st Jan 2000 is day 36526 etc.
- A *fractional* part of the day represents the *time of the day*. Lunch time on 1st April 1998 is 35886.50 (with 35886 representing the *Day* number corresponding to 1st April 1998 and the fractional .50 representing the exact middle of the day). 6:00 in the morning (1st April '98) is 35886.25 (one quarter of the day); at 8:00 in the morning is 35886.333333 (one third of the day) etc.

Internally, Excel doesn't handle dates (or time) any differently than any other number. The *only* difference lies in the *display* of the number: dates (and time) have a special *number (display) format*.

Normally, you don't have to be aware of the internal date arithmetic since you can enter a date (or time) directly into a cell and Excel will handle all the technicalities (i.e. convert it into its date number and apply the proper date format).

For instance, to enter the date *1st April, 1998* in a cell, you could enter any of the following:

1 April 1998 or 1-Apr-1998 or 1998/04/01 or 1998-04-01

To find the difference between two dates, just subtract the one from the other. The following spreadsheet calculates my age expressed as the number of days.

	A	B	C
1	Today:	08-Mar-98	=NOW()
2	Date of birth:	29-Jul-61	1961/07/01
3	Age in days:	13372	=B1-B2

Note the use of the date formula `=NOW()` which returns the current date (and time). There are many ways of formatting dates and a wide variety of functions that allow you to do so-phisticated date arithmetic. E.g. `WEEKDAY` allows you to determine which day of the week, `DAY` finds the day-of-the-month (1-31) etc.

Typically, date arithmetic works with the *non-decimal (integer)* part of the date number and time arithmetic with the *decimal (fractional)* part of date numbers.

Some notes on spreadsheet design

Think before you act

Because Excel makes it so easy to start building a model straight away, many people start constructing models before they have thought about the detailed design and layout. They then spend much time trying to change the format or, even worse, find that their model isn't flexible enough to accommodate changes at a later stage. So, before you rush out and start building a slightly more complex model, take some time to think through the logic and layout first.

A correct model is more important than attractive formatting

It is much more important to have a model that produces the correct figures than to have one with all the colours and borders you can think of but which produces incorrect result. First try to build a model that *works* correctly; once you are happy with the results (did you double-check?) you may want to spruce it up to make it more attractive.

Conversely, don't be overly impressed with other people's extremely attractive spreadsheet models: check their logic and see if the figures make sense.

Double-check your model's formulas with common sense and Excel's audit tools.

Before relying on the results of your model, take some time to check if the values in your spreadsheet are of the right order of magnitude. Though you may miss some finer errors, many incorrect spreadsheet formulas produce figures that are way off the normal "ball-park figures". A good way to check a spreadsheet is by plugging in some historical data and see how well the spreadsheet numbers correspond to actual figures. (Note: the process of removing mistakes from models, which are actually nothing but a type of financial computer program, is called *debugging*.)

Whether you suspect mistakes, you are busy building your own spreadsheet or even if you are only wanting to check the logic of someone else's spreadsheet, you are well advised to make full use of Excel's advanced auditing tools.

When you edit a cell's formula by pressing the <F2> edit key, Excel will use a colour coding scheme to indicate which cells are being referenced in the formula. Each cell reference in the formula is displayed in a different colour and the values in the corresponding cells have a border of the same colour around them.

You can get a more visual impression of the logic of your spreadsheet by looking how cells are interrelated through formulas. Use the **Tools • Formula Auditing** menu options for this. You may wish to have the *Formula Auditing* toolbar permanently available if you anticipate quite a bit of cell "auditing".

	A	B	C	D	E	F
1	Jean-Paul's Excel-nt Supermarket					
2				VATRate:	14%	
4	Item	Quantity	Price	Total (excl)	VAT	Total (incl)
5	Apples	3.5	R 3.49	R 12.22	R 1.71	R 13.93
6	Bananas	1.25	R 2.45	R 3.06	R 0.43	R 3.49
7	Cherries	0.75	R 7.45	R 5.59	R 0.78	R 6.37
8			Total:	R 20.87	R 2.92	R 23.79

The above figure shows a "first generation" of precedent arrows (**Tools • Formula Auditing** menu option / toolbar button). The arrow shows how the value in **F8** is derived from cells F5 to F7.

	A	B	C	D	E	F
4	Item	Quantity	Price	Total (excl)	VAT	Total (incl)
5	Apples	3.5	R 3.49	← R 12.22	← R 1.71	← R 13.93
6	Bananas	1.25	R 2.45	← R 3.06	← R 0.43	← R 3.49
7	Cherries	0.75	R 7.45	← R 5.59	← R 0.78	← R 6.37
8			Total:	R 20.87	R 2.92	R 23.79

By selecting the same option (or clicking the same button) again, you can determine where the values in F5 to F7 come from. A third click will trace the dependencies a further step back.

	A	B	C	D	E	F
1	Jean-Paul's Excel-nt Supermarket					
2				VATRate:	14%	
4	Item	Quantity	Price	Total (excl)	VAT	Total (incl)
5	Apples	← 3.5	← R 3.49	← R 12.22	← R 1.71	← R 13.93
6	Bananas	← 1.25	← R 2.45	← R 3.06	← R 0.43	← R 3.49
7	Cherries	← 0.75	← R 7.45	← R 5.59	← R 0.78	← R 6.37
8			Total:	R 20.87	R 2.92	R 23.79

You can use the *Audit* tool also to check dependencies in a "forward" manner i.e. which cells are dependent on (or influenced by) the current value. You can also trace where the root of all the *error values* that appear all of a sudden, come from.

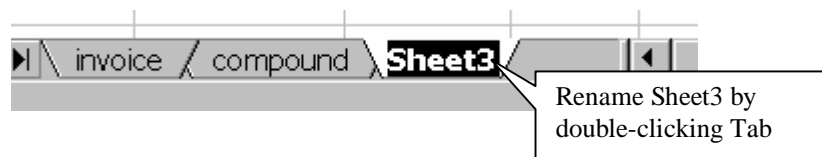
Use multiple sheets where applicable

If you have several spreadsheets that use the same basic data but are of a fundamentally different structure, keep them together (and linked) by using different sheets (or *pages*) in the same Excel document. This allows you to change the layout and formatting of any one sheet (e.g. changing column widths or inserting rows) without affecting the integrity of the data on the other sheets but it still keeps all data tightly grouped together.

RATE						
A	B	=vlookup(A5,list!A2:B6,2)				
1	Jean-Paul's Excel-nt Supermarket					
2				VATRate:	14%	
4	Item	Quantity	Price	Total (excl)	VAT	Total (incl)
5	Apples	3.5	A2:B6,2	R 12.22	R 1.71	R 13.93

Cell references to cells on other sheets become slightly more complicated since they now need to be prefaced with the name of the sheet and an exclamation mark ("!"). The above formula gives a simple example where the price of the item is stored on a sheet named *list* and can be looked up using the cell (range) reference *list!A9:B11*. (If the formula is to be copied, an absolute range reference should be used i.e. *list!\$A\$9:\$B\$11*.)

Note that a much better cell formula would be **=VLOOKUP(Item,PriceList,2)** with *Item* the "automatic" name for A5, and *PriceList* a name given to the range containing the price list. When using cell *names*, the reference to the sheet can be omitted and you also don't have to worry about the relative or absolute references any more. Yet another reason to use *range names*. You can rename a sheet by double-clicking (or right-clicking) the sheet tab at the left of the bottom scroll bar. (Cell ranges are named by selecting the range and clicking the *name box*.)



It is equally possible to make cell references to cells in other spreadsheets (Excel documents) but this practice is usually only recommended if documents become too large to be handled as separate work sheets.

Using different sheets is also recommended if your data has a distinct three-dimensional structure e.g. a large Excel document may contain sales data for many products (listed vertically, one row per product), for various regions (listed horizontally, one column for each region) and for a number of months or years (organised in different sheets, one sheet for each time period). Although Excel is configured initially to create new documents with a fixed number of sheets (usually three), you can easily add or delete sheets by right-clicking the sheet tabs or using the appropriate *Insert* or *Edit* menu options, respectively. You can also move sheets in a different sequence by clicking their tabs and dragging them across to their new position.

Note that it is possible to refer to cell ranges that span several sheets at once using a 3D reference e.g. **=SUM(Sheet1:Sheet3!A1:B2)** will sum the range A1:B2 across Sheet1, Sheet2 and Sheet3. You can also refer to a cell (or range) in another workbook e.g. **=SUM([Model.xls]Relations!D17:D24)** refers to the sheet "Relations" in the workbook file named "Model".

Creating Charts

Charts are an excellent way of presenting large amounts of numerical information in an easy and intuitive format. Excel makes it extremely easy to convert any data table into a chart quickly. On the other hand, Excel allows you to customise your chart to an amazing degree and you can spend many hours playing around with the many options and features which Excel offers.

When you create a chart, it becomes an intrinsic part of your spreadsheet; they either have their own *sheet* or they can appear anywhere inside one of the worksheets. Note also that charts always reflect the underlying data in the spreadsheet: as soon as the values change, the chart will be updated automatically.

The "quick and dirty" way to create a chart

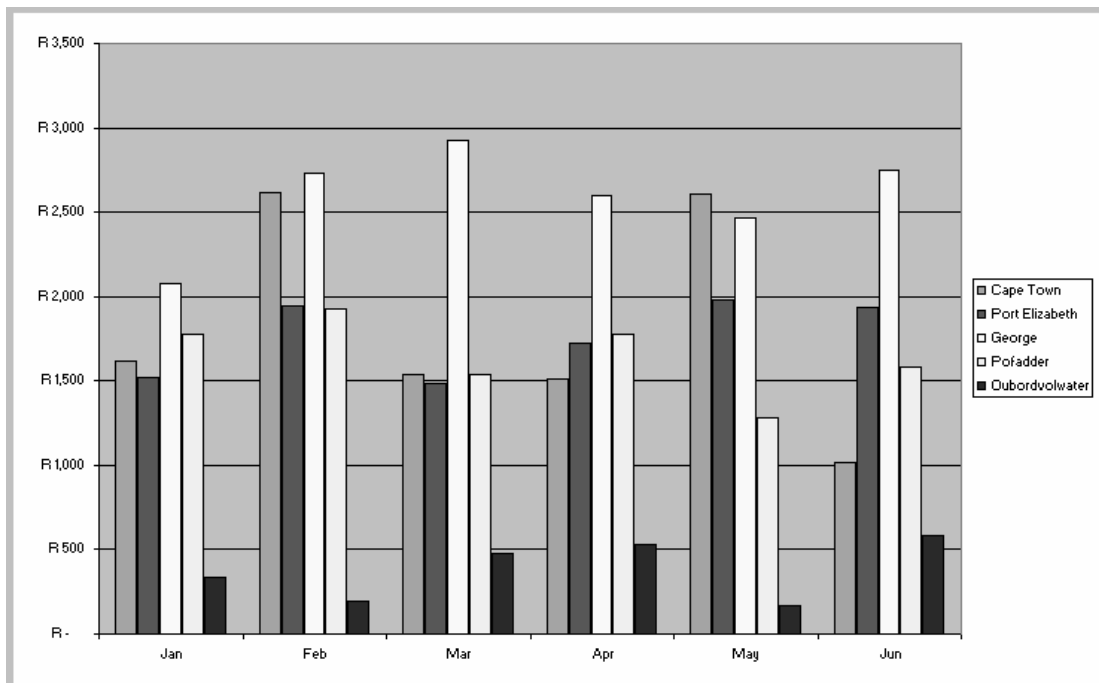
Before you create a chart, make sure that your data is organised in a table format. Excel usually takes a good guess about how your chart should look like, but it helps to give Excel a little hand by bearing the following in mind.

- Label your columns and rows with appropriate (and short) labels.
- Don't have a blank row/column of cells between labels and data.
- Preferably keep the data in contiguous cells.
- Sort your data columns and rows in the same order as you wish them to appear on the graph.

The following is an example of a table that is ready to be converted to a graph.

	A	B	C	D	E	F	G
1	<u>JP's Supermarket Sales</u>						
2							
3	Store	Jan	Feb	Mar	Apr	May	Jun
4	Cape Town	R 1,615	R 2,616	R 1,535	R 1,515	R 2,604	R 1,017
5	Port Elizabeth	R 1,521	R 1,941	R 1,487	R 1,721	R 1,979	R 1,940
6	George	R 2,078	R 2,728	R 2,926	R 2,598	R 2,469	R 2,748
7	Pofadder	R 1,779	R 1,925	R 1,538	R 1,779	R 1,280	R 1,583
8	Oubordvolwater	R 332	R 193	R 475	R 529	R 164	R 580

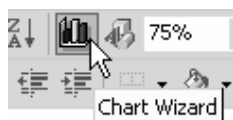
The quickest way to create a chart is by selecting any cell in the table and pressing the **<F11>** function key. This will produce a relatively unsophisticated but entirely acceptable chart. You have literally created a chart by just pressing one single key!



This chart will be put in a separate worksheet with the name **Chart1**. You will be able to customise this chart using the *Chart* command (which replaces the *Data* command on the menu whenever your cursor is on or in a chart).

Using the Chart Wizard.


If you want more control over the creation of your chart, use the *Chart Wizard* instead. Although the process takes slightly longer, you will save a lot of time because you won't have



to customise the chart afterwards. A *wizard* is an automated sequence of guided menus or windows that take you step-by-step through a process (in this case the creation of a chart). Position your cursor on a cell in the

table or, for even more control, select the area that contains the data to be charted. (Hint: you can select non-contiguous ranges by holding down the <Ctrl>-key when selecting non-adjacent ranges!)

First step of the Chart Wizard

Click the *Chart Wizard* button  on the standard toolbar (or select the **Insert • Chart...** menu option). The first screen will allow you to choose an appropriate chart type and sub-type.

The major chart types are the following.

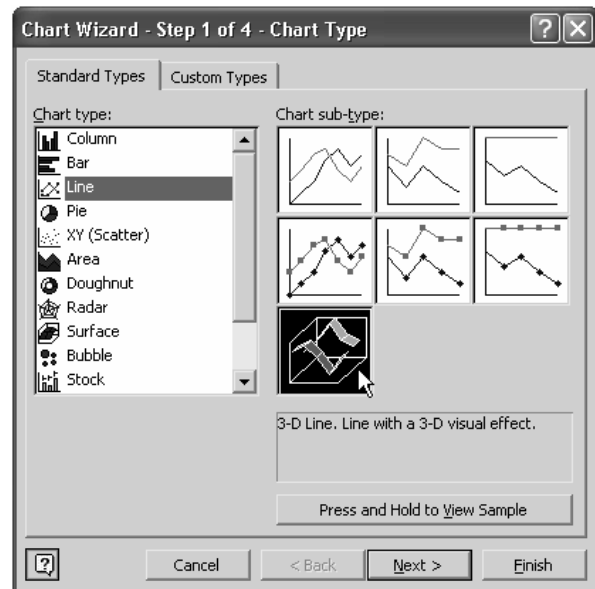
Column and Bar chart: stress the differences between various data points. Best used for categorical data (e.g. regions, product types etc.) but also acceptable for time series.

Line and Area chart: emphasise the *trend* displayed by the data. Use for continuous data such as time series, growth curves and other evolutionary processes.

Pie chart: shows the distribution of the data. The data series must add up to a total so that they can be expressed as a percentage.

XY scatter: for (2-dimensional) data points with no intrinsic sequence but where there is a specific relationship between the independent and dependent variable. Unlike the previous charts, you can have data points with identical independent values.

The *Chart sub-type* option allows you to refine your choice further. Usually the choice of sub-type depends on the relationship between the various dimensions of your data; e.g. some data can be represented as a stacked column chart (if the values are additive) but other data types cannot. Be careful with three-dimensional charts. Although they look more professional, they are often not so easy to read. On the other hand, they may enable you to visualise complex data much more easily.

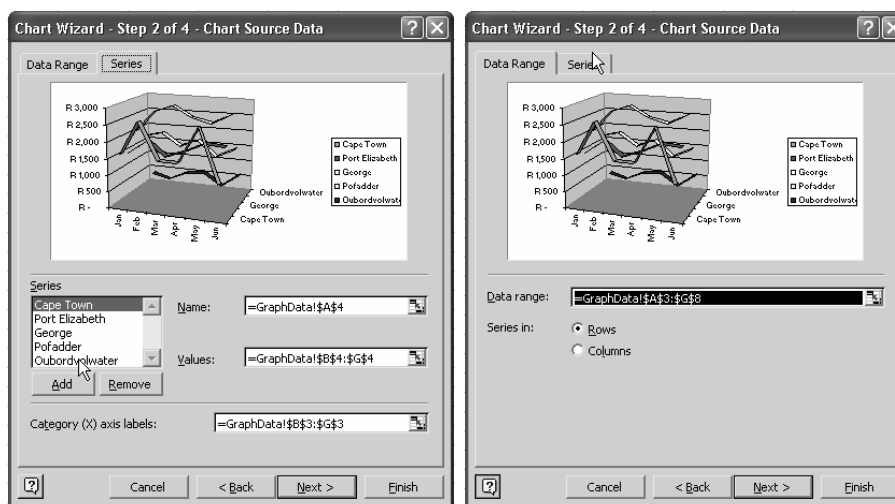


You can see a preview of how the chart (sub-)type will look for your own data by clicking and holding down the big *View Sample* button. When you are happy with your selection, click the *Next >* button.

Step Two of the Chart Wizard

This step allows you to select the orientation of the data. The *Data Range* tab allows you to indicate whether your data is organised horizontally (in rows) or in columns.

This is also the step where you have more detailed control over each of your data series. Click the *Series* Tab to delete or add data series, change titles for the series or edit the ranges where the series values are to be found.

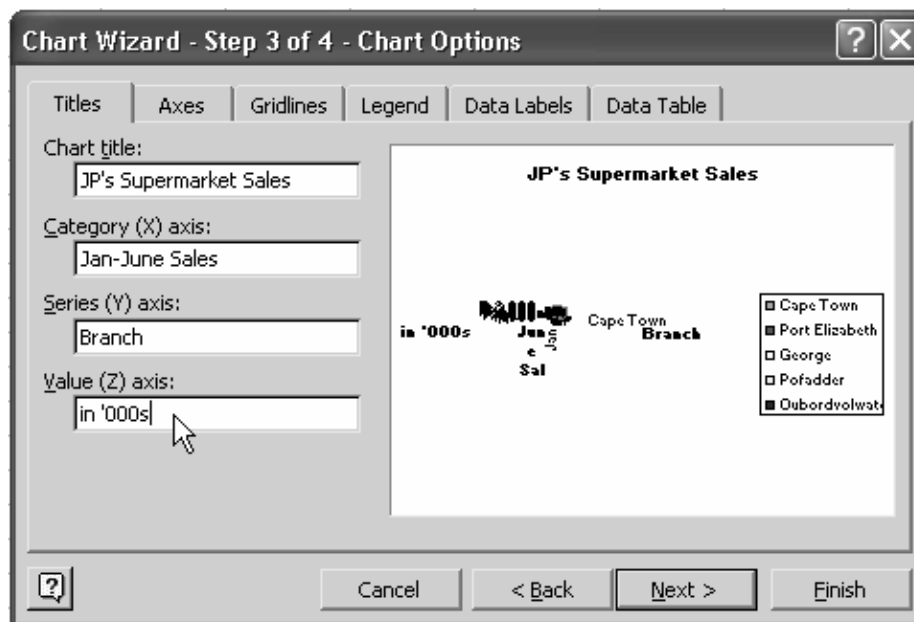


Remember that you can use the *collapse* icon to remove the window and keep only the text entry box visible for when you need to take a look at your spreadsheet (to check cell references).

Note that you may need to adjust the data range if your table includes column or row totals and you used the automatic range select prior to the invoking the graph wizard.

Step Three of the Chart Wizard.

Now that you have selected the *type* of chart and which *data* you want to chart, you are ready to fine-tune the exact formatting of the chart. There are many options available and we shall briefly highlight some of them.



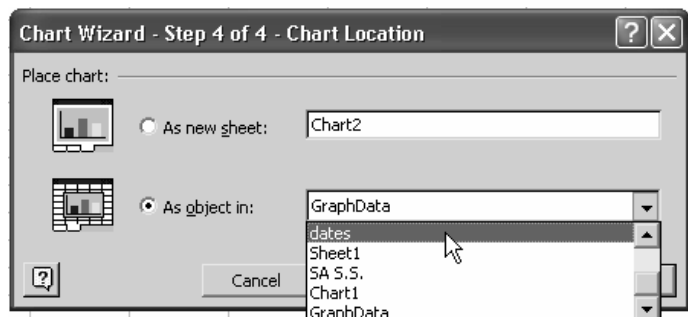
If you are unsure about the meaning of any of the following terms, refer to the finished chart on the next page.

- **Titles:** allows you to specify the title for the overall chart and label the various axes.
- **Axes:** you can set the format of the various axes and whether labels should be shown.
- **Gridlines:** specify whether grid lines show or not.
- **Legend:** where the legend should be positioned (or omit it altogether).
- **Data labels:** whether the actual values should be displayed next to the plotted data points (and where).
- **Data table:** allows you to include a formatted table of the plotted data as part of the chart.

Step Four of the Chart Wizard

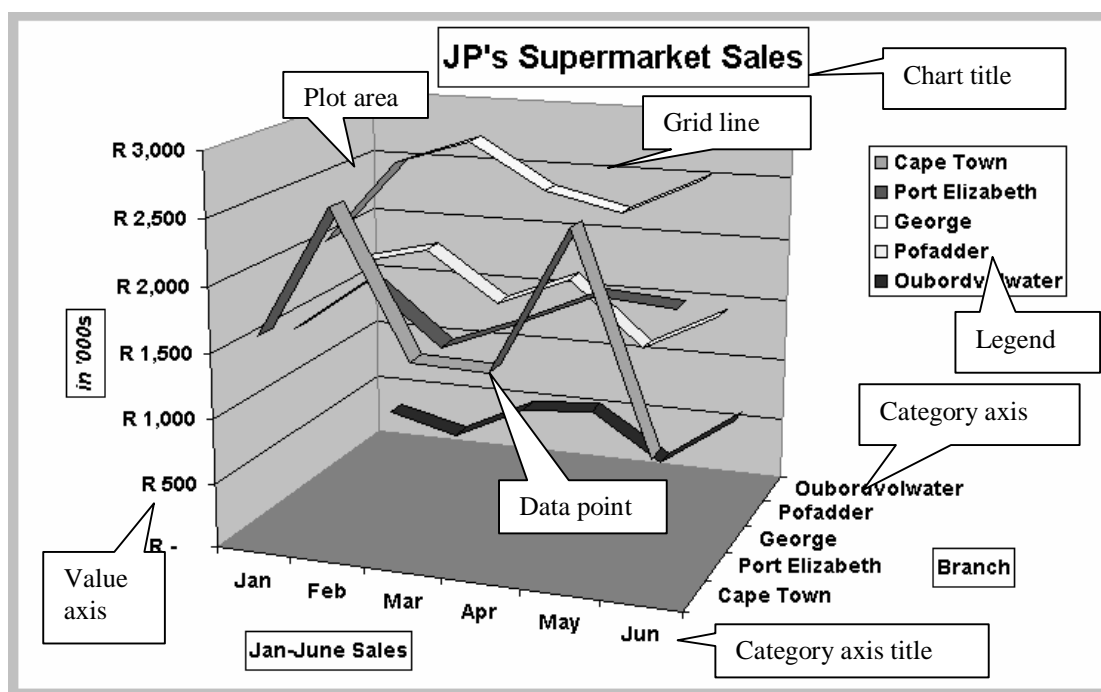
This final step controls where the chart should be placed in your document.

You can include it in your main worksheet (where it can be moved, resized and handled like any other embedded graphical *object*) or it can be put in its own separate chart sheet (no object formatting controls but easily accessible using the sheet tab).



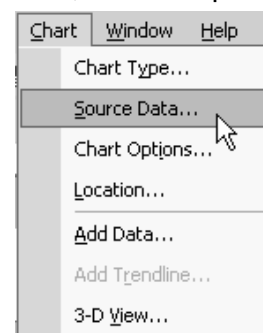
The various elements of a chart

A chart has a number of elements, each of which can be formatted and positioned separately.



The quickest way to customise any chart element is by right-clicking it. Or click it and press **<Ctrl>-<1>** to open up the formatting window. It is sometimes a little tricky to right-click the correct chart element e.g. the text box instead of the text inside the text box, or a data point instead of the chart wall, etc. But a little practice makes perfect.

You can also customise a chart by using the *Chart Toolbar* or the *Chart* menu options. The *Chart* menu replaces the *Data* menu whenever you are inside a chart.



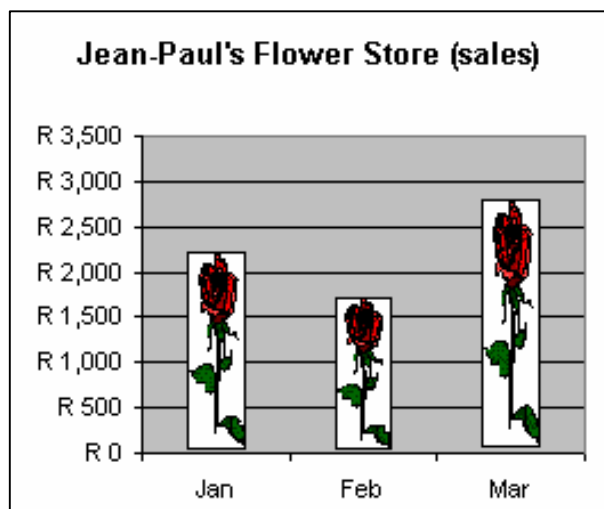
Further Graph Customisation

It is possible to further customise your chart e.g. by inserting WordArt, AutoShapes and other objects.



Creating a picture chart

You can make bar and column charts much more attractive by creating a picture chart. Copy the picture you wish to use (from the clip art library or graphics application) to the clipboard (*Edit • Copy* or **<Ctrl>-<C>**), select the data series (or data point) in Excel and select the *Edit • Paste* option (or **<Ctrl>-<V>**) to paste the picture into your chart.



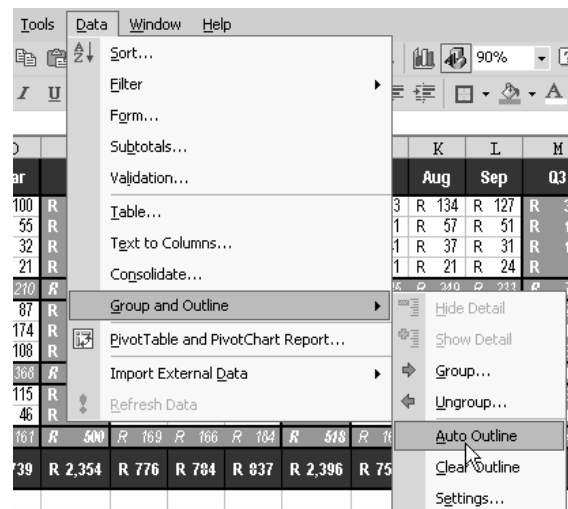
Grouping and Outlining Data Tables

If you have a fairly large, hierarchically structured data set with varying levels of aggregation which you want to explore at different levels of detail, you may want to use the grouping and outlining feature of Excel.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Sales (R mln.)	Jan	Feb	Mar	Q1	Apr	May	Jun	Q2	Jul	Aug	Sep	Q3	Oct	Nov	Dec	Q4	Year
2	Cape Town	R 138	R 112	R 100	R 350	R 103	R 104	R 108	R 315	R 113	R 134	R 127	R 374	R 103	R 126	R 122	R 351	R 1,390
3	Worcester	R 67	R 61	R 55	R 183	R 74	R 66	R 59	R 199	R 61	R 57	R 51	R 168	R 60	R 59	R 74	R 193	R 743
4	Paarl	R 34	R 32	R 32	R 99	R 38	R 34	R 33	R 105	R 41	R 37	R 31	R 109	R 40	R 31	R 32	R 103	R 417
5	Swellendam	R 29	R 22	R 21	R 72	R 20	R 29	R 28	R 78	R 21	R 21	R 24	R 65	R 29	R 28	R 29	R 86	R 302
6	Western Cape	R 268	R 228	R 210	R 706	R 236	R 234	R 228	R 697	R 236	R 249	R 233	R 717	R 233	R 243	R 267	R 733	R 2,852
7	Pretoria	R 95	R 105	R 87	R 287	R 92	R 106	R 115	R 313	R 106	R 82	R 81	R 268	R 92	R 114	R 114	R 320	R 1,189
8	Johannesburg	R 197	R 169	R 174	R 539	R 158	R 160	R 198	R 516	R 152	R 147	R 179	R 477	R 206	R 166	R 197	R 569	R 2,101
9	Soveto	R 92	R 123	R 108	R 323	R 122	R 118	R 113	R 353	R 103	R 98	R 123	R 324	R 122	R 119	R 91	R 331	R 1,330
10	Gauteng	R 383	R 397	R 368	R 1,149	R 372	R 385	R 425	R 1,182	R 361	R 326	R 382	R 1,069	R 420	R 399	R 402	R 1,220	R 4,620
11	Durban	R 120	R 109	R 115	R 343	R 118	R 121	R 131	R 370	R 106	R 131	R 114	R 351	R 117	R 99	R 93	R 309	R 1,374
12	Pietermaritzburg	R 55	R 56	R 46	R 157	R 51	R 44	R 52	R 148	R 56	R 42	R 45	R 143	R 53	R 47	R 52	R 151	R 598
13	KwaZulu-Natal	R 174	R 164	R 161	R 500	R 169	R 166	R 184	R 518	R 162	R 172	R 160	R 494	R 170	R 146	R 145	R 461	R 1,973
14	Total for SA	R 825	R 790	R 739	R 2,354	R 776	R 784	R 837	R 2,396	R 758	R 747	R 775	R 2,280	R 822	R 788	R 804	R 2,414	R 9,444


The table shown in the example, lists sales data horizontally across by month with subtotals for each quarter and the entire year. The rows represent individual stores, grouped by province.

Often you are only interested in the aggregated sub-totals. Select **Data • Group and Outline • Auto Outline** from the menu. Excel will now group the data at the same hierarchical levels based on the formulas used in the spreadsheet. Your spreadsheet will have collapse buttons next to the row / column names, which function similar to the folder collapse buttons in Explorer.

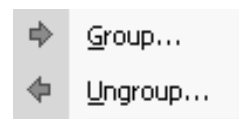


The image shows the Excel spreadsheet after applying the 'Auto Outline' feature. The data is now grouped, and collapse buttons are visible. A callout box points to the collapse buttons on the left, stating 'Hide all up to this level'. Another callout box points to the collapse buttons on the top, with labels 'Reveal', 'Range', and 'Hide'. A third callout box points to the collapse button next to the 'Sales (R mln.)' header, stating 'Lowest level'.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	Sales (R mln.)	Jan	Feb	Mar	Q1	Apr	May	Jun	Q2	Q3	Q4	Year						
2	Cape Town	R 138	R 112	R 100	R 350	R 103	R 104	R 108	R 315	R 374	R 351	R 1,390						
3	Worcester	R 67	R 61	R 55	R 183	R 74	R 66	R 59	R 199	R 168	R 193	R 743						
4	Paarl	R 34	R 32	R 32	R 99	R 38	R 34	R 33	R 105	R 109	R 103	R 417						
5	Swellendam	R 29	R 22	R 21	R 72	R 20	R 29	R 28	R 78	R 65	R 86	R 302						
6	Western Cape	R 268	R 228	R 210	R 706	R 236	R 234	R 228	R 697	R 717	R 733	R 2,852						
10	Gauteng	R 383	R 397	R 368	R 1,149	R 372	R 385	R 425	R 1,182	R 1,069	R 1,220	R 4,620						
13	KwaZulu-Natal	R 174	R 164	R 161	R 500	R 169	R 166	R 184	R 518	R 494	R 461	R 1,973						
14	Total for SA	R 825	R 790	R 739	R 2,354	R 776	R 784	R 837	R 2,396	R 2,280	R 2,414	R 9,444						

- Clicking on the **+** reveals the detail columns (or rows) for this sub-area. In the figure above, clicking the **+** button above the Q-column (where the cursor is currently positioned) will cause the columns containing the months of Q4 (the fourth quarter) i.e. columns N, O and P (Oct – Nov) to be displayed.
- Clicking on a **-** button hides the columns or rows as indicated by the associated bar: .
- Clicking on the level number indicators **1 2 3** causes all corresponding data rows/columns to be collapsed or expanded simultaneously. E.g. click on the **3** to display the lowest level of detail (i.e. by month or store). Click on **2** to hide the lowest level and see only the quarterly or provincial subtotals. Clicking on **1** will display only the highest level i.e. the total for the year or the whole of South Africa.

Since the the AutoOutline feature is based on the logic of the formulas used in your spread sheet, it may not always generate the correct levels. It is possible to manually adjust the grouping. Select a range containing the column or rows you wish to group (or ungroup) and click the **Group** (or **Ungroup**) menu option on the **Data • Group and Outline** menu (or the respective buttons on the toolbar if you've added the buttons.)



Select the **Rows** or **Columns** option in the *Group* window and the rows or columns within the highlighted range will now be grouped accordingly. If you already had several levels of group/outline, a new level will be added.

Simple Data Analysis Using Data Functions

Excel is often used to capture and process simple data contained in table format. Examples of simple data needs are address lists, pricing tables, sales transactions, and student marks. Indeed, a famous predecessor of Excel was called Lotus 1-2-3 with the “1” referring to the (financial) formula modelling capabilities, the “2” to its graphics capabilities and the “3” to its data capabilities.

Data that can be stored in a single table is called “**flat data**” or “**list data**”, as opposed to more complex data as contained in hierarchical, network or relational format. Because a separate module deals with the use of Access for processing complex data, only a rudimentary overview of the list data analysis capabilities of Excel will be given here. Generally, each **row** contains a “**record**” of data. For the examples listed above these would be (in respective order: a person, a product item, a sales transaction or a student. **Columns** would hold data **attributes** or **fields**. For the address list, fields could be first name, surname, telephone number, street address, birthday etc. For a pricing table, different columns would hold the product code, product name, product category, cost price, mark-up percentage and list sales price.

It is suggested (and hereby assumed) that the first row of a data table, the **header row**, contains the *names* (or descriptions) of the data **fields**. The row immediately below the header row should **not** be left blank but be filled with the data for the first record (item). It is bad practice to include *any* blank rows in a data table –they may cause the data analysis functions to work in an unpredictable fashion. In particular, the first column of a data table should contain a continuous set of data all the way down to the last data record. Ideally, the first column contains an identifier which is *unique* to each row such as an ID number or even a row number. Conveniently, when you have calculated fields (fields which are based on other fields using a formula, e.g. Net Sales = Sales – Discount) in your table, Excel will automatically copy them when entering (adding) a new line at the bottom of your data table.

To illustrate the data analysis function, a table with 998 raw sales transactions is used. The data fields are:

- Transaction number (Xaction): a sequential number from 39283 to 40280
- Region: the regional area where the sales was made; a choice of WestCape, Gauteng, KwazuluN and NorthCape.
- Product: which product item was sold i.e. Gizmos, Thingies, Widgets or Gadgets.
- Month: the month in which this

	A	B	C	D	E	F	G
1	Sales Report						
2	Xaction	Region	Product	Month	CustType	Discount	Sales
3	39283	Gauteng	Gizmos	Apr	Corporate	10%	R 125.94
4	39284	WestCape	Gizmos	Apr	Wholesale	25%	R 1,206.24
5	39285	KwazuluN	Widgets	Mar	Wholesale	10%	R 245.42
6	39286	WestCape	Gadgets	May	Corporate	0%	R 517.25
7	39287	WestCape	Widgets	Jan	Wholesale	0%	R 595.25
8	39288	Gauteng	Widgets	Apr	Corporate	10%	R 2,095.54
9	39289	WestCape	Thingies	Jun	Corporate	25%	R 1,327.73
10	39290	Gauteng	Thingies	May	Export	25%	R 2,216.10
11	39291	Gauteng	Thingies	Jan	Corporate	0%	R 621.31
12	39292	WestCape	Gadgets	Mar	Corporate	0%	R 710.77
13	39293	KwazuluN	Thingies	May	Corporate	10%	R 514.38
14	39294	WestCape	Gizmos	Jun	Wholesale	10%	R 803.78
15	39295	Gauteng	Gadgets	Mar	Individual	25%	R 1,058.59
16	39296	WestCape	Gizmos	Apr	Wholesale	10%	R 118.16
17	39297	NorthCape	Gizmos	Mar	Individual	0%	R 461.50
18	39298	WestCape	Thingies	Apr	Corporate	0%	R 552.15
19	39299	NorthCape	Gadgets	Jun	Individual	0%	R 168.49
20	39300	WestCape	Widgets	May	Wholesale	10%	R 1,921.61
21	39301	WestCape	Widgets	May	Wholesale	25%	R 1,593.26
22	39302	WestCape	Thingies	Apr	Wholesale	25%	R 333.31
23	39303	KwazuluN	Gadgets	Feb	Corporate	25%	R 727.32
24	39304	NorthCape	Thingies	May	Export	0%	R 925.22

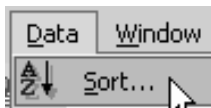
item was sold.

- CustType: the type of customer (Corporate, Individual, Export or Wholesale).
- Discount: the discount percentage applied to the sale (0, 10 or 25%).
- Sales: the actual sales amount.

Note that data assumed to be irrelevant for high-level data analysis, e.g. the customer name, day of the month or actual store was already filtered out.

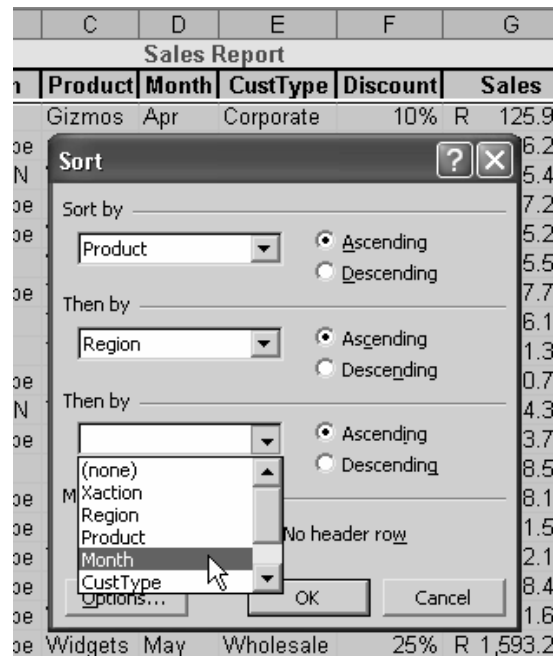
Sorting Data

The most frequently used data tool is sorting data in some sort of alphabetical order. If your data is fully continuous, you can click anywhere within your table, else you should select the range of the entire table manually.



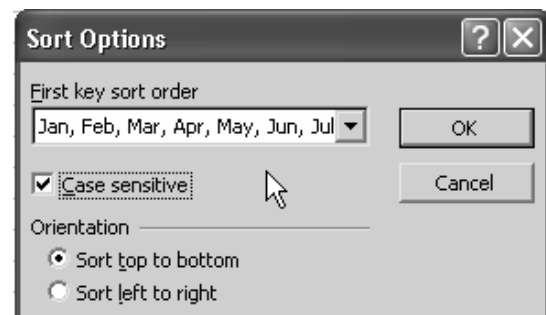
Select the **Sort...** option on the **Data** menu and this will present you with the *Sort* dialogue window.

You can choose up to three fields to sort your data with, with a different sorting order for each: **ascending** i.e. from A • Z for text data or low to high for numbers; or **descending** i.e. high to low. If your data table contains a header row, you will be able to select the sorting key (i.e. the field) based on the field names; else you will be presented with the column names.



If all you need is a quick, simple sort of your data table on just one particular field, you can just position your cursor in the field you want the table to be sorted on and click one of the appropriate toolbar buttons for an ascending sort respectively. For a multi-level or **nested** sort you need to select the **Data • Sort** option. In the example above, the entire table will first be sorted according to *Product*. The rows containing the same product will be sub-sorted according to *Region*. The rows with the same product and region will then be sorted according to *Month*.

Note that the month used here is a textual string, so the following order will be used: April < Feb < Jan < June < Mar < May! If a *true* date was used i.e. a date number formatted in a particular data number format, there would be no problem. The advanced options actually allow you to specify a non-normal sort (for the first key sort). In addition, you can also change the orientation of the sort (horizontal instead of vertical) and make the sort case sensitive, i.e. a < b < z < A < B.

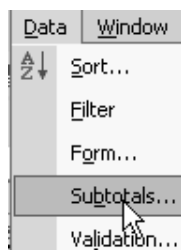


Be warned: once data is sorted and the file has been saved, it cannot be unsorted again. Many data sets have become irreversibly corrupted by accidentally selecting only a few of the fields of the data set and sorting those fields whilst the unselected fields remained in the original order. In a large dataset, this often goes undetected until it is too late to *undo*... For this and other reasons, it is good practical to insert an extra column containing the original row order numbers. (In our example, this is not necessary because the transaction numbers can fulfil the same function since they are ordered and unique.)

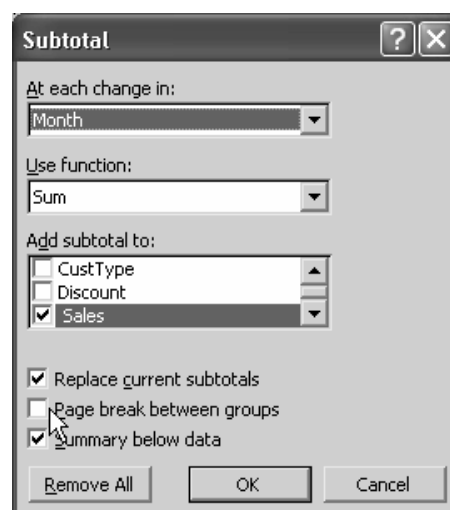
Automatically Creating Subtotals

If your data contains a number of categories, you can create subtotals for each of your categories.

The subtotal command assumes that you have sorted your data according to the category, e.g. if we want the monthly subtotal for each region and each product, then we would sort according to (1) region, (2) product and then (3) month as illustrated above. If monthly totals are required across all regions and products, then re-sort the data using the *month* as the *first* sort key.



To create (or remove) subtotals, select the **Subtotals...** option on the **Data** menu. This brings up the *Subtotal* dialogue window which allows you to specify the category for which you want to create subtotals (the table should have been sorted on this category), what mathematical function should be used to create the subtotal and for which fields you wish to create the subtotals.



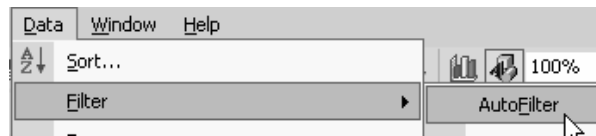
The *Page break between groups* option is especially useful for when you intend printing a report with each category on a separate page.

	A	B	C	D	E	F	G
1	Sales Report						
2	Xaction	Region	Product	Month	CustType	Discount	Sales
111	40012	KwazuluN	Gadgets	Feb	Corporate	25%	R 1,635.31
112	40224	KwazuluN	Gadgets	Feb	Corporate	10%	R 13.60
113	40274	KwazuluN	Gadgets	Feb	Corporate	0%	R 700.25
114	Feb Total						R 8,417.62
115	39371	KwazuluN	Gadgets	Jan	Export	0%	R 1.86
116	39535	KwazuluN	Gadgets	Jan	Wholesale	10%	R 178.07
117	39646	KwazuluN	Gadgets	Jan	Export	25%	R 572.63
118	39679	KwazuluN	Gadgets	Jan	Export	25%	R 950.46
119	39822	KwazuluN	Gadgets	Jan	Export	10%	R 177.48
120	39862	KwazuluN	Gadgets	Jan	Export	25%	R 1,620.60
121	40099	KwazuluN	Gadgets	Jan	Corporate	25%	R 672.48
122	40124	KwazuluN	Gadgets	Jan	Wholesale	25%	R 1,607.47
123	Jan Total						R 5,781.05
124	39352	KwazuluN	Gadgets	Jun	Export	0%	R 264.24
125	39361	KwazuluN	Gadgets	Jun	Wholesale	10%	R 988.57
126	39370	KwazuluN	Gadgets	Jun	Corporate	10%	R 608.02
127	39380	KwazuluN	Gadgets	Jun	Corporate	25%	R 1,110.05

Apart from the *Sum*-function, you may find that you use the *Count* and *Average* functions also quite frequently for tallying the subtotals. Note also that Excel conveniently creates outline buttons allowing you to quickly switch between the detailed and the subtotal view.

Data Filter

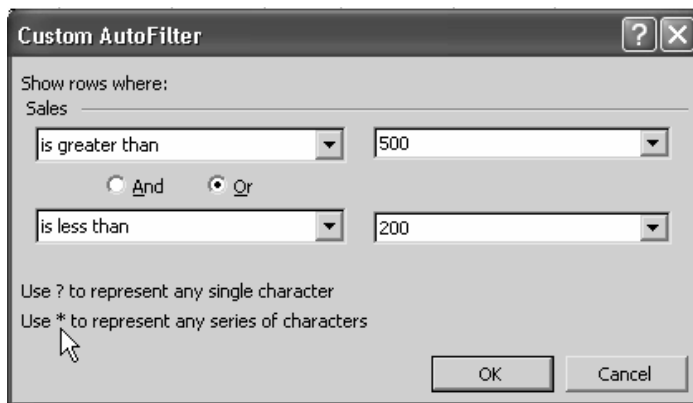
Another very useful and simple function for interactive data analysis is the **Data • Filter** menu option. This allows you to display or hide selected rows, based on the value of one or several fields. The most commonly used filter with *categorical* data is the **AutoFilter** option.



	A	B	C	D	E	F	G
1	Sales Report						
2	Xacti	Region	Produ	Mon	CustTyp	Discou	Sales
9	3928	(All)	Thingies	Jun	Corporate	25%	R 1,327.73
29	3930	(Top 10...)	Thingies	Jun	Corporate	0%	R 984.30
39	3931	(Custom...)	Thingies	Jun	Individual	25%	R 972.92
51	3933	Gauteng	Thingies	Jun	Individual	0%	R 446.83
56	3933	KwazuluN	Thingies	Jun	Wholesale	10%	R 465.29
62	3934	NorthCape	Thingies	Jun	Export	10%	R 1,775.92
86	39366	WestCape	Thingies	Jun	Individual	10%	R 297.62
98	39378	KwazuluN	Thingies	Jun	Wholesale	10%	R 155.68
113	39393	KwazuluN	Thingies	Jun	Corporate	10%	R 904.90

Clicking this option will create drop-down lists of filter-values for each of the fields. Click on any drop-down list button to only display the desired categories. E.g. in the example below, only the sales

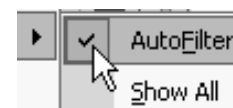
for “Product = Thingies” and “Month = Jun” are being displayed (the list-button changes colour for those fields where a filter is active). The screen shows how an additional filter to select only the “Region = Gauteng” is being selected. Note that all filter selections are applied simultaneously (i.e. “and” instead of “or”).



For more complex selections or for specifying criteria on non-categorical data (e.g. values), use the **(Custom...)** option on the drop down list. This allows you to specify more specific numerical filter criteria or to include *wildcard* characters (? or *) in text filters.

down lists altogether, select the **Data • Filter • AutoFilter** menu

To remove the autofilter drop-



option again. Autofilter is a toggle option and the little tick indicates whether the option is turn on (in effect) or off.



The **Advanced Filter...** option on the **Data • Filter** menu is used to specify more complex criteria e.g. specific required value combinations of fields or multiple allowable values. One particularly useful application of this is where a long list contains duplicate records and you wish to create a (separate) subset containing only unique records.

Advanced Data Analysis using Pivot Tables

What Are Pivot Tables?

A more advanced way of doing data analysis is by means of *pivot tables*. This is an extremely powerful tool for interactive data exploration and **data mining**: looking for hidden relationships or patterns in large volumes of data. Again, it is most useful if you have a number of different categories or groups according to which data can be classified. Each category (field) constitutes a potential **dimension** for analysis. In data mining your data set is called the **data cube**. If all your dimensions are numerical - continuous values - instead of categorical - a limited range of discrete values - you should use Excel's statistical analysis tools instead e.g. multiple regression or ANOVA.

The distinction between numerical and categorical data is quite important for pivot tables.

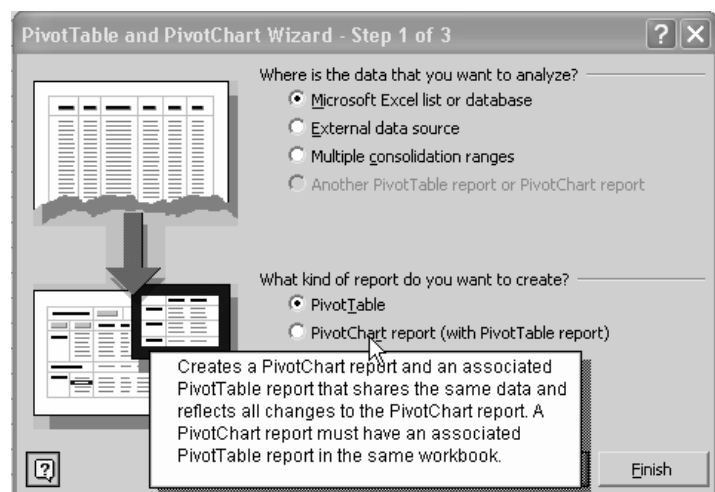
- **Numerical** data consists of values, probably unique to each record. Examples are sales, costs, marks, scores and other values. This type of data can be totalled, averaged, and you can calculate the highest/lowest value, standard deviation etc.
- **Categorical** data consists of a limited set of options. Examples are sales region (only a certain number of provinces), customer type (e.g. corporate, private or wholesale), product group etc. These can be only tabulated (listed), usually in a sorted order, or the number of records within a given category can be counted.

The pivot table will interactively tabulate subtotals (or other statistics) for a chosen numerical dimension (e.g. sales) according to any selected categorical dimensions (according to customer type and sales region). It is basically an interactive, explorative, multi-dimensional, table-oriented and very flexible version of the subtotal function explained in a previous subsection.

Although seemingly complex at first, Excel's pivot table feature is actually not all that difficult to use thanks to the excellent pivot table wizard. Note that the exact use of the wizard and the pivot table feature has changed (improved) quite substantially in the last few Excel upgrades, so if you are using an earlier version of Excel, some of the details may differ. If you are a student, your instructor may provide you with a good (i.e. relatively large) data set to practice with. We will illustrate the pivot table using the sales report example from the previous section.

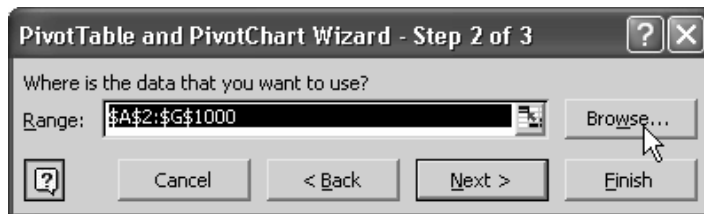
Using the Pivot Table Wizard

To start, click anywhere on your dataset (again – we assume that it is a continuous dataset i.e. without blank columns or rows, and a proper header row indicating the field names). Select the ***PivotTable and PivotChart Report...*** option





from the **Data** menu. This brings up the first step of the *PivotTable* Wizard which allows you to specify your source and required output. Your data source is most likely to be a *Microsoft Excel list or database*, although you may want to use an external (flat) database table accessible through the Office Data Connection. This is useful if you have a very large set of data e.g. more than 65535 records (the maximum an Excel sheet can handle). The output report can be the PivotTable with or without the PivotChart.



The second step of the PivotTable wizard allows you to specify the range of your dataset (the interface will be more complicated if you use a non-Excel source). If your dataset is in another spreadsheet, you can

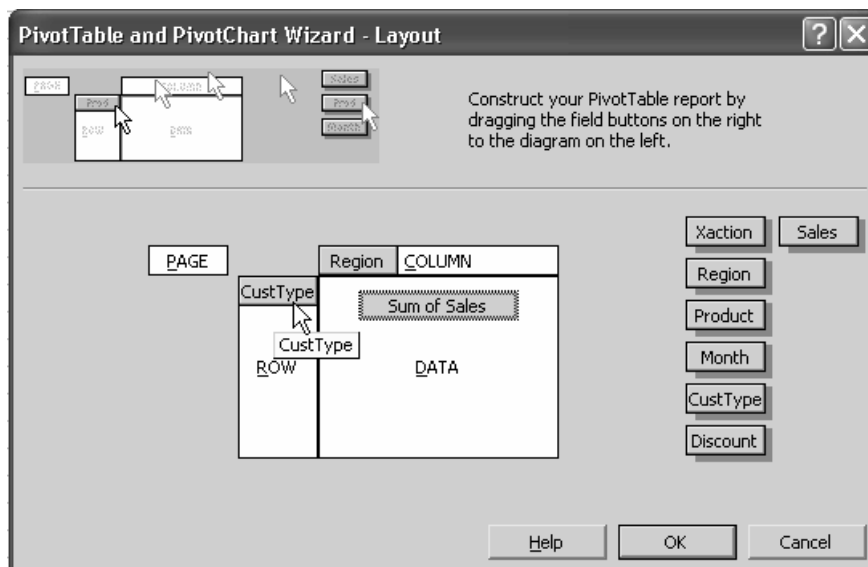
click the **Browse...** button to choose the source file.

The last step of the wizard allows you to specify whether you want the report in a **New** or an **Existing** worksheet. It is preferable to put the report in a separate i.e. new worksheet.



An additional option in this last step is the PivotTable

Options..., which you are unlikely to change in your first few sessions.



Another, more important, option is to specify an initial layout for your pivot table. Click the **Layout...** button to access the *PivotTable Layout* dialogue window. On the right-hand side, you will see the list of your fields. In the middle you will find a blank pivot table report layout.

The easiest decision is what is the actual (numerical) data that you wish to analyse, should be. In this case, it is the actual *sales* figures. Drag the **Sales** button from the right and drop it into the **DATA** area of the pivot table report. Excel knows that it is numerical data and will assume that you want to analyse the *Sum of Sales*, i.e. subtotals. (You can change the type

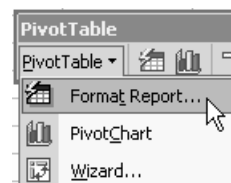
of function easily enough in the report at a later stage). Next you need to select two initial *categorical* fields you want to use to construct your table, i.e. two dimensions which you suspect of showing some pattern. In the example, we investigate to see if there is any special relationship between the type of customer (**CustType**) and the sales **Region**. Each button has been dragged across and dropped on the **ROW** and **COLUMN** area respectively.

Using the Pivot Table

Click **OK** to set your initial layout and click **Finish** to complete the last step of the *pivot table* wizard. The result will be a screen similar to the one below.

Sum of Sales	Region				
CustType	Gauteng	KwazuluN	NorthCape	WestCape	Grand Total
Corporate	R 76,659	R 69,168	R 42,181	R 46,914	R 234,922
Export	R 67,830	R 43,856	R 49,205	R 44,325	R 205,217
Individual	R 49,627	R 27,988	R 34,749	R 35,993	R 148,356
Wholesale	R 67,358	R 47,381	R 43,441	R 55,062	R 213,243
Grand Total	R 261,474	R 188,394	R 169,571	R 181,737	

Note the actual Pivot Table report, the pop-up task pane containing the field names, and the Pivot Table toolbar. Click the **Format Report...** option on the **PivotTable** menu on the *PivotTable* toolbar to apply an AutoFormat to your pivot table or apply your own formatting using the standard formatting options. (We will not go into detail on how to format the report.)



We will also not endeavour to actually analyse any of the above data but of course, this is actually the purpose of the entire pivot table analysis. Generally, you should look for exceptionally large or small values, e.g. you can note that the lowest sales are the sales to individuals in the Kwazulu-Natal region. Also, the Northern Cape seems to be exporting relatively more (because of its proximity to Namibia? Botswana?) than e.g. Kwazulu-Natal or the Western Cape, despite the fact that its overall sales are lower!

Assume that we want to do a more in-depth analysis of the sales in the Cape. Click on the drop down field name for the rows, **Region**, and deselect the *Gauteng* and *KwazuluN* boxes so that only the Northern and Western Cape remain. Click **OK** and a new table, a subset of the previous one will appear. We are interested particularly in comparing the *Export* and the *Wholesale*, since the sales for the other customer types appears to be normal. So deselect the *Corporate* and *Individual* options from the **CustType** dropdown list to have an even more reduced table.

Sum of Sales	Region
Corporate	
Export	
Individual	
Wholesale	
Grand Total	

Let us now explore in more depth the remaining categories and bring in a third dimension. Is there any effect from the *discount*? Drag the **Discount** field button from the *Field List* task pane across to drop it on the column heading *CustType*, in cell A4 (or use the **Add To** button to add it to the *Column Area*). The pivot table will now include the *Discount* dimension.


	A	B	C	D	E	F	G	H
1	Drop Page Fields Here					PivotTable Field List		
2						Drag items to the PivotTable report		
3	Sum of Sales		Region			<ul style="list-style-type: none"> [-] Xaction [-] Region [-] Product [-] Month [-] CustType [-] Discount [-] Sales 		
4	CustType	Discount	NorthCape	WestCape	Grand Total	<input type="button" value="Add To"/> Row Area		
5	Export	0%	R 6,893	R 9,014	R 15,907			
6		10%	R 14,544	R 16,043	R 30,588			
7		25%	R 27,768	R 19,268	R 47,036			
8	Export Total		R 49,205	R 44,325	R 93,530			
9	Wholesale	0%	R 16,261	R 11,319	R 27,581			
10		10%	R 13,056	R 15,154	R 28,209			
11		25%	R 14,124	R 28,589	R 42,713			
12	Wholesale Total		R 43,441	R 55,062	R 98,503			
13	Grand Total		R 92,646	R 99,388	R 192,034			
14								
15								

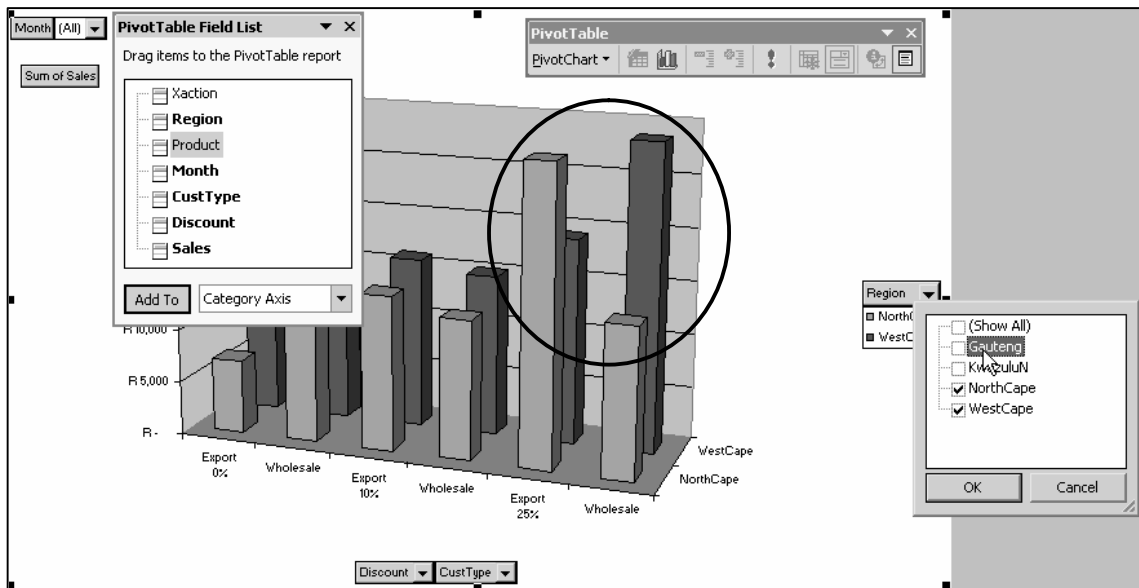
It is clear in the table that the Northern Cape has been promoting its exports by awarding big discounts (25%) while the Western Cape has been pushing its wholesale sales by awarding the 25% discounts. This can be visualised even better by swapping the **Discount** and the **CustType** buttons: drag the **Discount** button (B4) left and drop it on **CustType** (A4) to produce the differently organised pivot table below.

	A	B	C	D	E
1	Drop Page Fields Here				
2					
3	Sum of Sales		Region		
4	Discount	CustType	NorthCape	WestCape	Grand Total
5	0%	Export	R 6,893	R 9,014	R 15,907
6		Wholesale	R 16,261	R 11,319	R 27,581
7	0% Total		R 23,155	R 20,333	R 43,488
8	10%	Export	R 14,544	R 16,043	R 30,588
9		Wholesale	R 13,056	R 15,154	R 28,209
10	10% Total		R 27,600	R 31,197	R 58,797
11	25%	Export	R 27,768	R 19,268	R 47,036
12		Wholesale	R 14,124	R 28,589	R 42,713
13	25% Total		R 41,891	R 47,857	R 89,749
14	Grand Total		R 92,646	R 99,388	R 192,034


The top row is generally used as a filter mechanism, by dropping Field Buttons in the top row. E.g. the *Month* field button was dragged from the task pane to row 1. Then the month of **Mar** was selected from the dropdown list button (in B1) in order to display only the sales for March.

	A	B	C	D	E	F	G	H
1	Month	(All)						
2								
3	Sum of Sales			WestCape	Grand Total			
4	Discount			9,014	R 15,907			
5	0%			11,319	R 27,581			
6				20,333	R 43,488			
7	0% Total			16,043	R 30,588			
8	10%			15,154	R 28,209			
9				31,197	R 58,797			
10	10% Total			19,268	R 47,036			
11	25%	Export						
12		Wholesale	R 14,124	R 28,589	R 42,713			
13	25% Total		R 41,891	R 47,857	R 89,749			
14	Grand Total		R 92,646	R 99,388	R 192,034			
15								

For people who are more comfortable with a more visual approach, you should generate the Pivot Table Chart e.g. by clicking the Pivot Chart button  on the toolbar. This chart has similar options and functions to the Pivot Table Report, but is often a lot easier to interpret. Below is the previous report as a 3-D chart (you may need to change the chart type to suit



your style and data needs). Note how the Export-Wholesale reversal is highly visible in the chart area.

There are many more exciting options. For instance, you may want to change the formula used to summarize the data. Use the pivot table *Field Settings* button  or menu option to e.g. display the *average* instead of the *sum* of the sales. You can even express them as % of subtotals (click on the **Options** button).

In the end, the only real way to learn to use the pivot table feature is by practicing and exploring it.

Exercises

Exercise 1

Build the spreadsheet as requested below.

Sheet 1

This sheet compares the (fictional) airfares to popular destinations by a number of airlines.

	A	B	C	D	E	F	G	H
1	<i>To:</i>	<i>London</i>	<i>New York</i>	<i>Toronto</i>	<i>Hawai</i>	<i>Average fare</i>	<i>Rating (Text)</i>	<i>Rating (Graphic)</i>
2	<i>SAA</i>	R 4,300	<i>R 5,800</i>	<i>R 5,950</i>	<i>R 6,800</i>	<i>R 5,713</i>	extreme	ⓂⓂ
3	<i>BA</i>	R 3,400	R 4,900	R 5,000	<i>R 6,650</i>	R 4,988	medium	Ⓜ
4	<i>KLM</i>	R 3,950	<i>R 5,450</i>	<i>R 5,600</i>	<i>R 6,800</i>	<i>R 5,450</i>	expensive	Ⓜ
5	<i>AirEgypt</i>	R 2,590	R 3,800	R 3,950	<i>R 6,800</i>	R 4,285	reasonable	Ⓜ
6	<i>Virgin</i>	R 2,480	R 3,400	R 4,100	<i>R 5,990</i>	R 3,993	cheap	ⓂⓂ
7	middle fare:	R 3,400	R 4,900	R 5,000	<i>R 6,800</i>			
8	range:	R 1,820	R 2,400	R 2,000	R 810			
9								
10	Average Fare	Type	Symbol					
11	R 0	cheap	ⓂⓂ					
12	R 4,000	reasonable	Ⓜ					
13	R 4,500	medium	Ⓜ					
14	R 5,000	expensive	Ⓜ					
15	R 5,500	extreme	ⓂⓂ					

- Calculate the following:
 - Average fare = the average of the fares for each airline.
 - Rating = rating type for each airline based on their average fare, as per *rating table* (A11:C15)
 - Both a text and a graphical rating are given. The graphics are created using the Wingding font with the capital letters J, K & L.
 - Middle fare = the fare of the airline right in the middle of the range of fares e.g. BA's fare to London falls inbetween (or is the median value between) SAA & KLM's higher fares and AirEgypt & Virgin's lower fares to London.
 - Range = difference between the highest (=maximum) and lowest airfare to each destination.
- AutoFormat the table using the Colourful 1 format. You may have to re-format row 1 and columns A, H.
- Use a conditional format to show all the airfares above R5,000 automatically in italic bold
- Headings in row 1 should be set to automatically wrap around if their length exceeds cell width, and are to be aligned to the top of the cells.
- Cell A1 should have a diagonal line through it as per example (don't worry about the text alignment).

Sheet 2

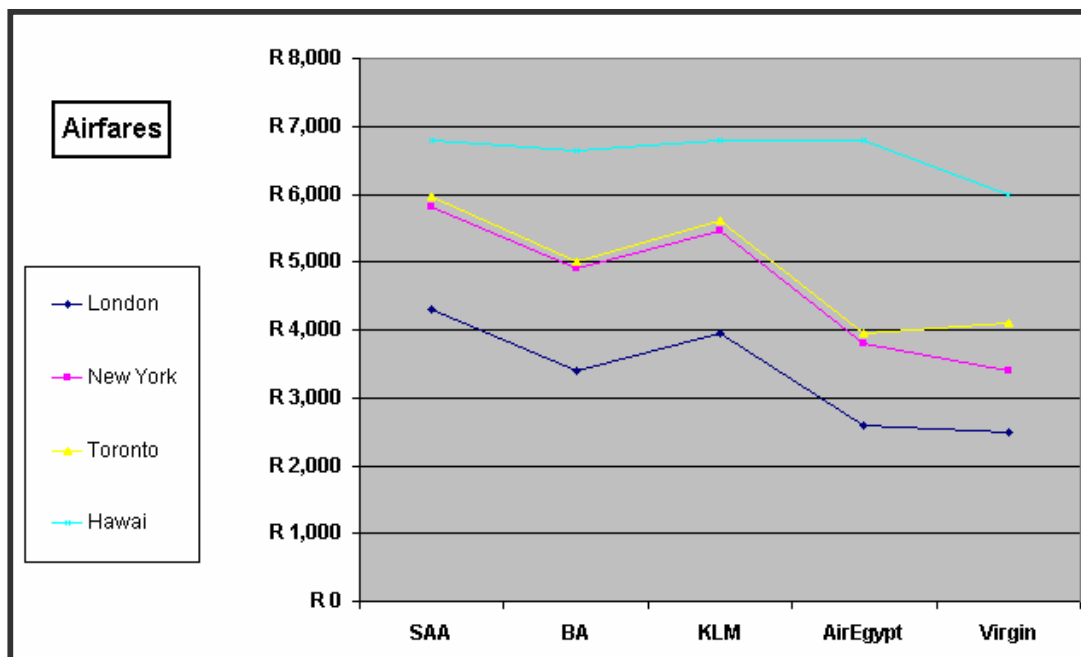
Open Sheet2 in your workbook (i.e. in the same spreadsheet document)

	A	B	C	D	C	D	E	F	G	H	I	J	K	L	M	N	O
1																	
2	Date	Day	Nr	x	Nr	x	5	6	7	8	9	10	11	12	13	14	
3	31-Aug-98	Mon	2	11	2	11	55	66	77	88	99	110	121	132	143	154	
4	01-Sep-98	Tue	4	12	4	12	60	72	84	96	108	120	132	144	156	168	
5	02-Sep-98	Thu	6	13	6	13	65	78	91	104	117	130	143	156	169	182	
6	03-Sep-98	Fri	8	14	8	14	70	84	98	112	126	140	154	168	182	196	
7	04-Sep-98	Mon	2	15	2	15	75	90	105	120	135	150	165	180	195	210	
8	05-Sep-98	Tue	3	16	3	16	80	96	112	128	144	160	176	192	208	224	
9	06-Sep-98	Thu	7	17	7	17	85	102	119	136	153	170	187	204	221	238	
10	07-Sep-98	Fri	1	18	1	18	90	108	126	144	162	180	198	216	234	252	
11	08-Sep-98	Mon	2	19	2	19	95	114	133	152	171	190	209	228	247	266	
12	09-Sep-98	Tue	8	20	8	20	100	120	140	160	180	200	220	240	260	280	
13	10-Sep-98	Thu	3		3												

- Use AutoFill to fill the range A3:A102 with the consecutive days starting from 31-Aug-98.
- Use AutoFill to fill the range B3:B102 with the cycle of the weekdays (excluding Wed) i.e. Mon, Tue, Thu, Fri, Mon, Tue, Thu, Fri, Mon etc.
- Use a formula based on a built-in function to fill the range C3:C102 with random integer numbers from 0 to (but never equal to) 10 (no decimals). The numbers should change each time the spreadsheet is updated. (2 marks!)
- Ensure that all values in the range C3:C102 are displayed sideways (i.e. at an angle of 90 degrees) and with a double underline.
- In D2:N12, create a table that calculates the product of the numbers [11 to 20] and [5 to 14] i.e. a “times” table as per example. You are required to use a formula that contains a mixed cell reference. (2 marks!)
- Split the window vertically as per example.
- Rename Sheet2 to Times and Sheet1 to Fares
- Add a new sheet with the name NewSheet and move it so that it sits between Times and Sheet3.
- Use the statistical built-in Excel formula to calculate the one-tailed probability of a value of 4.5 for the chi-square distribution. (5 degrees of freedom.) Answer: %
(Note: give the answer in %)

Chart

- Use the chart wizard to create a line chart representing the fares for each destination.
- Use a “Line-with-markers” chart
- Use the data range, titles, axes labels as per example.
- Give the chart the title Airfares which should have a bold border around it and be positioned on the left.
- Ensure that the legend is on the left hand side
- Put your chart on a separate Chart Sheet with the name Graph.
- Increase the titles, axes and legend fonts size to 14 or 16 points.



Exercise 2

Build the spreadsheet as requested below.

Sheet 1

We are tracking the sales for 6 branches and 3 products. Enter and format the data as per example below.

	A	B	C	D	E	F	G	H	I
1	Branch	Model1	Model2	Model3	Total	Budget	Variance	Verdict	Bonus?
2	Bloemfn	14	8	21	43	40	7.0%	Ok	Watch
3	Cape Town	30	10	20	60	55	8.3%	Ok	Hotel week-end
4	Durban	28	14	15	57	60	-5.3%	Not Ok!	nothing
5	George	10	5	7	22	25	-13.6%	Not Ok!	nothing
6	Joburg	40	18	13	71	70	1.4%	Ok	Pen
7	Windhoek	18	11	19	48	40	16.7%	Ok	Overseas' trip
8	Total:	140	66	95	301	290	14%		
9	Highest:	40	18	21	71	70	17%		
10	Average:	23	11	16	50	48	2%		
11	Exceeds budget by:	Bonus Incentive:							
12		0 Pen							
13		2 Watch							
14		4 Hotel week-end							
15		8 Overseas' trip							

- You should use the Excel built-in formulas to calculate the following:
 - Total (column E) = the sum total for the three models.
 - Budget = given (i.e. constants, no formulas)
 - Variance = the difference between total actual sales and the budget, expressed as a percentage of the budget. Note: this is the “accounting variance”, not the statistical one (i.e. don’t use Excel’s built-in *var* formula for this!)
 - Verdict = “Ok” if the variance is not negative, “Not Ok!” if the variance is negative.
 - Bonus = Bonus given for exceeding the budget. The applicable incentive bonus can be found in the “**bonus**” table A12:B15. Where the budget has *not* been achieved (i.e. variance is negative), *nothing* should be given as bonus. (If you cannot do the later without modifying the bonus table, a “#NA” is fine).
 - Total (row 8) = total for all branches
 - Highest = highest value for all branches
 - Average = average value for all branches
- AutoFormat** the table using the *list 3* format.
- Assign the name **Bonus** to the range A12:B15

- The headings of the bonus table should be set to *wrap*.
- The headings in row 1 should be at an angle of 45 degrees in 9 pt size.

All values should be centered, left-aligned and formatted to the required precision as per example.

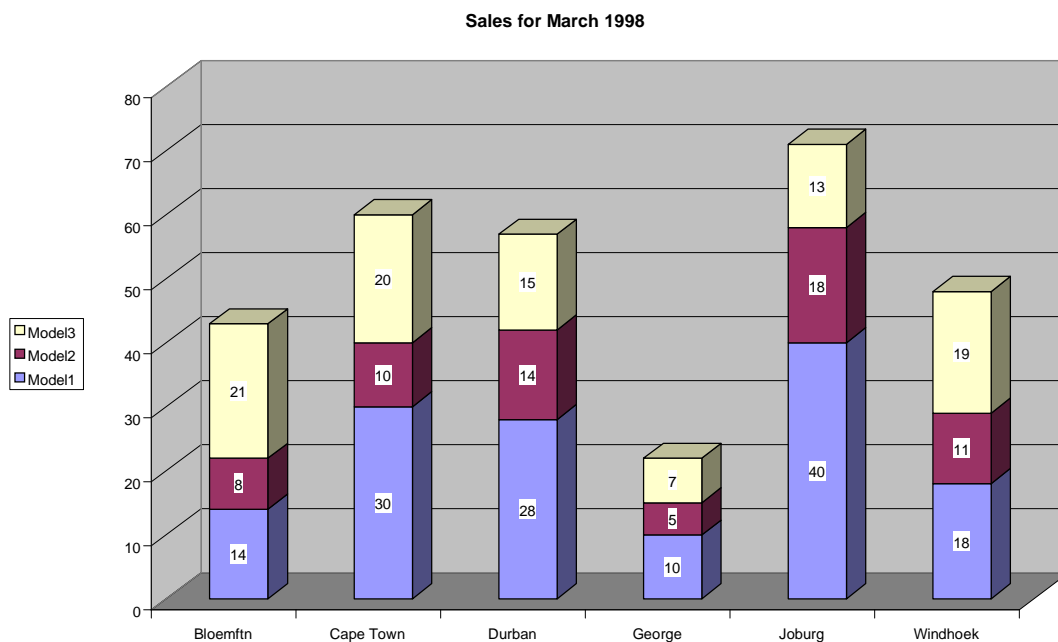
Sheet 2

- Open *Sheet2* in your workbook (i.e. in the *same* spreadsheet document)
- Use *AutoFill* to fill the range A1:A50 with consecutive dates:
- Use *AutoFill* to fill the range B1:B50 with the number sequence *1st, 2nd, 3rd, 4th, 5th etc.*
- Use *AutoFill* to fill the range C1:B50 with the sequence *Jan, , ...*
- Ensure that all values in the range B1:B50 are centered and *double* underlined.
- In D2:J11, create a “divided by” table which divides the numbers 100 to 1400 by selected *odd* numbers as per example. You are required to use a formula that contains a *mixed cell reference*.
- Split the window vertically.
- Rename *sheet2* to *Divide*
- Save your spreadsheet again (overwrite the old copy when you are sure that your data is still ok.)
- Delete *sheet3* in your workbook.

	A	B	C	D	E	D	E	F	G	H	I	J
1	01-Apr-98	<u>1st</u>	Jan			"Divided by" table						
2	02-Apr-98	<u>2nd</u>	Jun	/	3	/	3	7	11	13	17	
3	03-Apr-98	<u>3rd</u>	Nov	100	33.33	100	33.33	14.29	9.09	7.69	5.88	
4	04-Apr-98	<u>4th</u>	Apr	200	66.67	200	66.67	28.57	18.18	15.38	11.76	
5	05-Apr-98	<u>5th</u>	Sep	300	100.00	300	100.00	42.86	27.27	23.08	17.65	
6	06-Apr-98	<u>6th</u>	Feb	400	133.33	400	133.33	57.14	36.36	30.77	23.53	
7	07-Apr-98	<u>7th</u>	Jul	500	166.67	500	166.67	71.43	45.45	38.46	29.41	
8	08-Apr-98	<u>8th</u>	Dec	600	200.00	600	200.00	85.71	54.55	46.15	35.29	
9	09-Apr-98	<u>9th</u>	May	700	233.33	700	233.33	100.00	63.64	53.85	41.18	
10	10-Apr-98	<u>10th</u>	Oct	800	266.67	800	266.67	114.29	72.73	61.54	47.06	
11	11-Apr-98	<u>11th</u>	Mar	900	300.00	900	300.00	128.57	81.82	69.23	52.94	
12	12-Apr-98	<u>12th</u>	Aug	1000	333.33	1000	333.33	142.86	90.91	76.92	58.82	
13	13-Apr-98	<u>13th</u>	Jan	1100	366.67	1100	366.67	157.14	100.00	84.62	64.71	
14	14-Apr-98	<u>14th</u>	Jun	1200	400.00	1200	400.00	171.43	109.09	92.31	70.59	
15	15-Apr-98	<u>15th</u>	Nov	1300	433.33	1300	433.33	185.71	118.18	100.00	76.47	
16	16-Apr-98	<u>16th</u>	Apr	1400	466.67	1400	466.67	200.00	127.27	107.69	82.35	
17	17-Apr-98	<u>17th</u>	Sen									

Chart

- Use the chart wizard to create a chart representing the sales for the three models and the five branches.
- Use a *3-D Stacked Column* chart
- Use the data range, titles, axes labels as per example
- The legend should be on the left.
- Make sure that the actual (data) values are also shown in the columns
- Create the chart on a new sheet with the name *Sales Chart*
- Save your spreadsheet



Exercise 3

Build the spreadsheet as requested below.

Sheet 1

We have a small class of four students.

- Enter their names and marks for tests 1 to 3 as per example below. The maximum score for each test is 10. We award symbols as per the *symbol table* in A9:B15.
- You should use the Excel built-in formulas to calculate the following:
 - Total = the sum total of the marks for tests 1 to 3 (i.e. out of a possible maximum of 30)
 - Final = the total expressed as a percentage and *rounded to the nearest percent*.
 - Pass/Fail = Pass if the student obtains at least 50%, Fail if the student obtains less than 50%
 - Symbol = the appropriate symbol listed in the symbol table.
 - Average = average score for the class.
 - Highest = highest score obtained by any student in the class.
- **AutoFormat** the table using the *colorful 1* format.
- Assign the name **Symbols** to the range A11:B15
- Center the word **Symbol Table** across cells A9 and A10
- The headings *Less than* and *symbol* in row 10 should be at an angle of 45 degrees in 8 pt size.

Save your spreadsheet.

	A	B	C	D	E	F	G	H	I
1	Student	Test1	Test2	Test3	Total	Final	Pass/Fail	Symbol	
2	John	1.0	5.0	7.0	13.0	43%	Fail	D	
3	Joe	6.0	3.0	6.0	15.0	50%	Pass	C	
4	Jill	8.0	5.0	7.5	20.5	68%	Pass	B	
5	Jane	10.0	4.5	9.0	23.5	78%	Pass	A	
6	Average:	6.3	4.4	7.4	18.0	60%			
7	Highest:	10.0	5.0	9.0	23.5	78%			
8									
9	Symbol Table								
10	From:	Symbol:							
11	0%	E							
12	40%	D							
13	50%	C							
14	60%	B							
15	70%	A							

Sheet 2

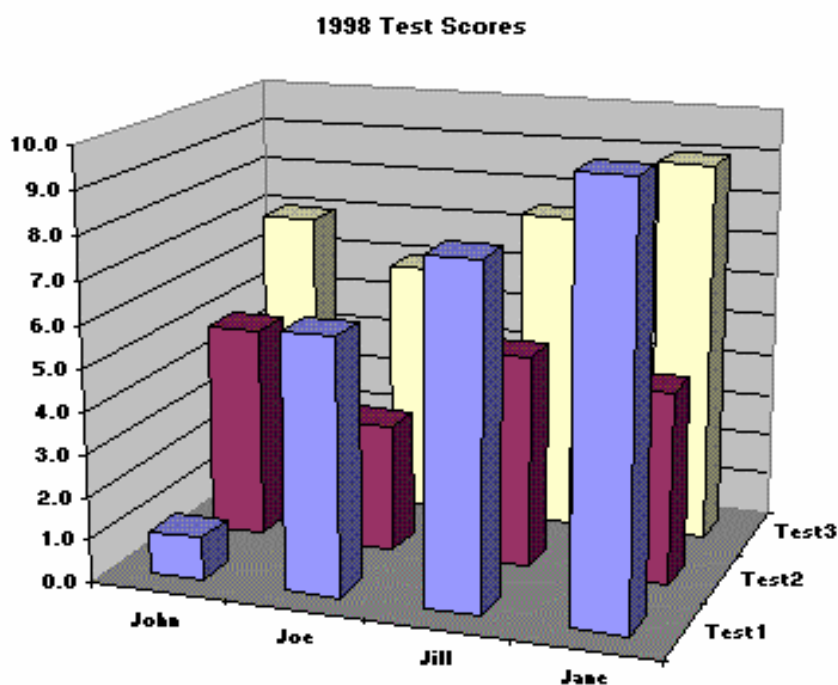
- Open *Sheet2* in your workbook (i.e. in the *same* spreadsheet document)
- Use *AutoFill* to fill the range A1:A50 with the days of the week: *Monday, Tuesday, ...*
- Use *AutoFill* to fill the range B1:B50 with the number sequence *15, 20, 25, 30, 35, 40, ...*
- Use *AutoFill* to fill the range C1:B50 with the sequence *Test 1, Test 2, Test 3, Test 4, ...*
- Ensure that all values in the range B1:B50 are centered and *double* underlined.
- In cell D13, use a built-in Excel formula to calculate the *AVERAGE* of the values in the range B1:B50.
- In cell D14, use a built-in Excel formula to calculate the *HIGHEST* value in the range B10 to B30
- Format cells D13 and D14 so that the values are displayed in a *FRAC-TION* format with *tenths* (one digit).
- In D2:J11, create a “times” table for the numbers (5 to 10) x (5 to 13) as per example. You are required to use a formula that contains a *mixed cell reference*.
- Format all columns so that their width is automatically set to fit the widest entry.
- Split the window both horizontally and vertically.
- Rename *sheet2* to *AutoFill*
- Save your spreadsheet again (overwrite the old copy when you are sure that your data is still ok.)
- Delete *sheet3*

	D
13	137 1/2
14	160

	A	B	C	D	E	F	G	H	I	J	
1	Monday	<u>15</u>	Test 1	Times Table							
2	Tuesday	<u>20</u>	Test 2	x	5	6	7	8	9	10	
3	Wednesday	<u>25</u>	Test 3	5	25	30	35	40	45	50	
4	Thursday	<u>30</u>	Test 4	6	30	36	42	48	54	60	
5	Friday	<u>35</u>	Test 5	7	35	42	49	56	63	70	
6	Saturday	<u>40</u>	Test 6	8	40	48	56	64	72	80	
7	Sunday	<u>45</u>	Test 7	9	45	54	63	72	81	90	
8	Monday	<u>50</u>	Test 8	10	50	60	70	80	90	100	
9	Tuesday	<u>55</u>	Test 9	11	55	66	77	88	99	110	
10	Wednesday	<u>60</u>	Test 10	12	60	72	84	96	108	120	
11	Thursday	<u>65</u>	Test 11	13	65	78	91	104	117	130	
12	Friday	<u>70</u>	Test 12								
47	Friday	<u>245</u>	Test 47								
48	Saturday	<u>250</u>	Test 48								
49	Sunday	<u>255</u>	Test 49								
50	Monday	<u>260</u>	Test 50								
51											

Chart

- Use the chart wizard to create a chart representing the scores for the four students and the three tests.
- Use a *3-D Column* chart
- Use the data range, titles, axes labels as per example
- There is *no* need for a legend - remove it.
- Give the chart the title *1998 Test Scores*
- Create the chart on a new sheet with the name *Scores Chart*
- Save your spreadsheet



Objectives 1

Efficient Data Entry 2

Using AutoFill 2

Copying the contents of a cell or range to adjacent cells. 2

Using AutoFill for common data series. 2

Using AutoFill to enter a linear series 3

AutoComplete 4

Other Automatic Data Entry Tools 4

More Advanced Cell Formatting 5

The Formatting Toolbar 5

Selecting more advanced cell formatting 6

Font Formats 6

Text Alignment within Cells 7

Borders, Colours, Shading and Patterns 8

Row and Column Formatting 8

Changing the width of a column 8

Hide / Unhide Columns 9

Additional notes 9

Changing the height of rows 9

AutoFormatting a range 10

Conditional Formatting 11

Number Formatting 12

Common number formats 12

What you see is *NOT* what you get 13

Automatic Number Formatting 14

Excel Formulas 15

Excel's Built-in Functions: the Basics 15

Statistical functions. 17

Mathematical functions 17

The Excel IF function 19

Lookup and reference functions 20

Convert Calculated Values in Constants Using Paste Special 24

More on Using Logical Conditions 24

Mixed addressing 25

Using Headers as Cell Range Names 26

Print Options 28

The Print dialogue window 28

Changing the Page Layout 28

Setting Sheet Options 29

Adjusting Print Margins 29

Viewing large spreadsheets 30

Working with dates and time 31

Some notes on spreadsheet design 32

Think before you act 32

A correct model is more important than attractive formatting 32

Double-check your model's formulas with common sense and Excel's audit tools. 32

Use multiple sheets where applicable 33

Creating Charts 35

The "quick and dirty" way to create a chart 35

Using the Chart Wizard. 36

The various elements of a chart 39

Further Graph Customisation 40

Creating a picture chart 40

Grouping and Outlining Data Tables 41

Simple Data Analysis Using Data Functions 43

Sorting Data 44

Automatically Creating Subtotals 45

Data Filter 46

Advanced Data Analysis using Pivot Tables 47

What Are Pivot Tables? 47

Using the Pivot Table Wizard 47

Using the Pivot Table 49

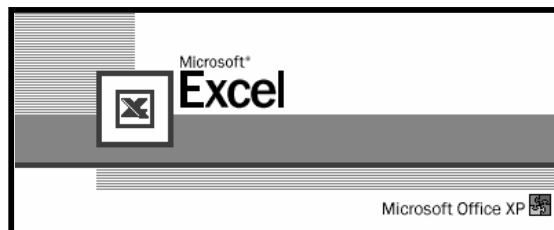
Exercises 52

Exercise 1 52

Exercise 2 55

Exercise 3 58

Using Excel 2002



© 1998-2005, Jean-Paul Van Belle