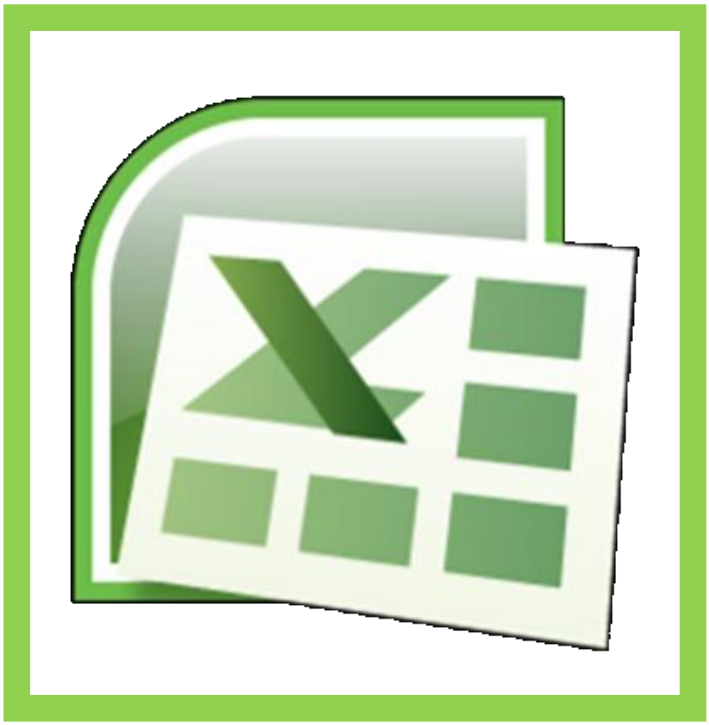
IGCSE ICT – SECTION 14
DATA ANALYSIS

MICROSOFT EXCEL
STEP BY STEP GUIDE



Mark Nicholls

ICT Lounge

Data Analysis – Self Study Guide

Contents

In this section you will learn how to:

- Enter text and number data into a spreadsheet
- Use editing functions such as cut, copy and paste
- Enter formulae and simple functions into a spreadsheet
- Replicate formulae and functions in a spreadsheet
- Test the spreadsheet for functionality
- Select subsets of data within the spreadsheet
- Change display and formatting of cells within a spreadsheet
- Change size of rows and columns within a spreadsheet
- Adjust the page orientation
- Save a spreadsheet
- Print a spreadsheet showing formulae or values
- Create a graph or a chart
- Label a graph or a chart
- Change chart colours to print in black and white

For this section you will need these source files from your teacher:

- CLASSLIST.CSV
- SALARY.CSV
- COSTS.CSV
- TASKS.CSV
- JOBS.CSV
- TUTORS.CSV
- RAINFALL.CSV
- ROOMS.CSV
- CLUBS.CSV
- STAFF.CSV
- ITEMS.CSV
- TUCKSHOP.CSV
- PROJECT.CSV
- CLIENT.CSV
- SALES.CSV
- EMPLOYEES.CSV
- TEACHERS.CSV
- OPERATORS.CSV
- WEBHITS.CSV

14.1–What is a data model?

For data analysis you will use a **spreadsheet model** to explore different possible answers. Models are sometimes called a **'What if'** scenario.

Models let you change data in the spreadsheet to see what will happen to the results.

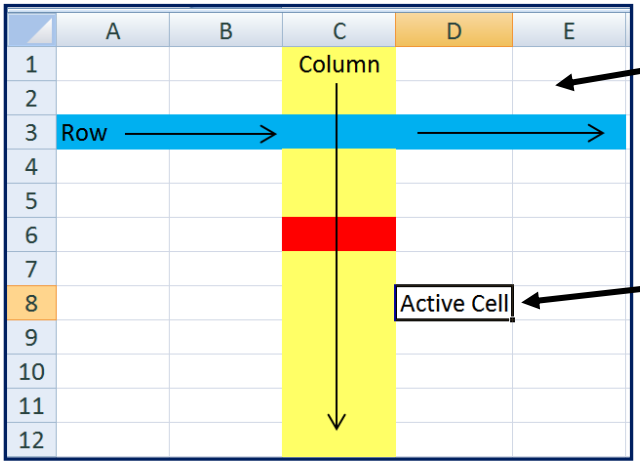
NOTE: In the practical examination you will be asked to build a simple spreadsheet model and make changes within it to produce different results.

14.1a – Spreadsheet Basics

You will use the spreadsheet software **Microsoft Excel** to create your data models.

Layout of a spreadsheet

A spreadsheet is a table which is split into **rows** and **columns**. The table is made up of a number of **cells**. It looks like this.



The Active Cell

The **Active Cell** is the cell which you have currently selected. It will have a darker outline around it so you can easily see which cell you are currently using.

Cell References

Each cell has a **unique address**. This address is known as the **'Cell Reference'** and it helps us identify cells for use in formulae. The cell reference comes from the **Column Letter followed by the Row Number**.

For example, the red cell in the picture above has a cell reference of **C6**. The active cell has a cell reference of **D8**.

Task A

Create a spreadsheet to **multiply** any **two numbers together** and display the result.

Save as **Data Analysis – Task A**

How to do it:

Open up **Microsoft Excel 2007**.

Each cell in a spreadsheet can hold one of three things:

- A **number**
- Text (often referred to as a **label**)
- A **formula** (which always starts with an = sign).

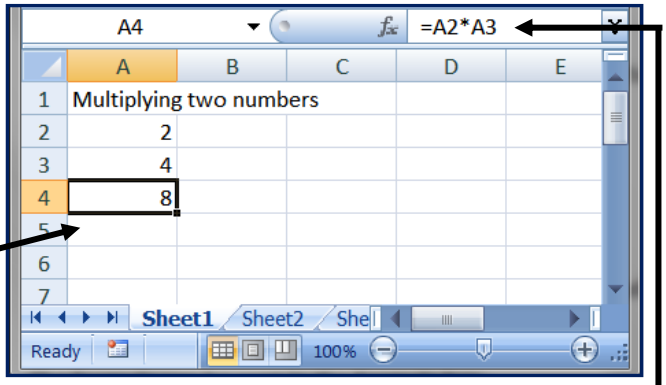
Move the cursor into cell **A1** and type in the label **'Multiplying two numbers'**.

Move the cursor into cell **A2** and enter a number. Repeat this for cell**A3**.

In cell **A4**, enter the following formula then press enter:

=A2*A3

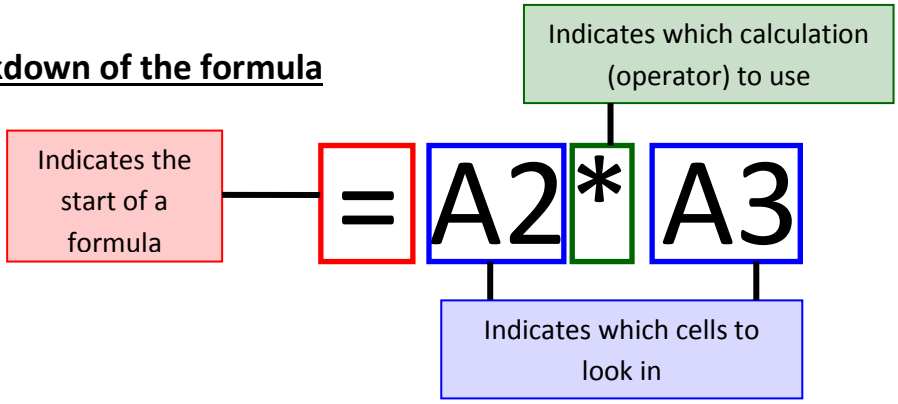
The spreadsheet should look like this:



Notice how the formula is not visible in the cell. The cell contains the result of the formula.

The formula can be seen in the **formula bar**.

Breakdown of the formula

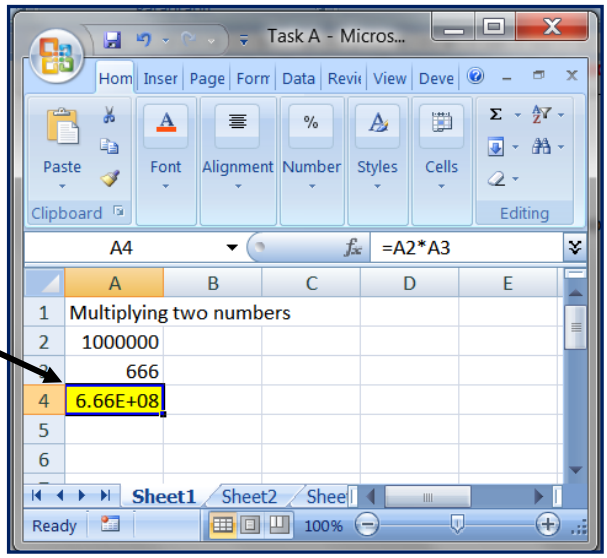


NOTE: If you created the spreadsheet as shown, you should be able to change the contents of cells A2 and A3 to multiply any two numbers together.

The changing of cells to see new results is called modelling.

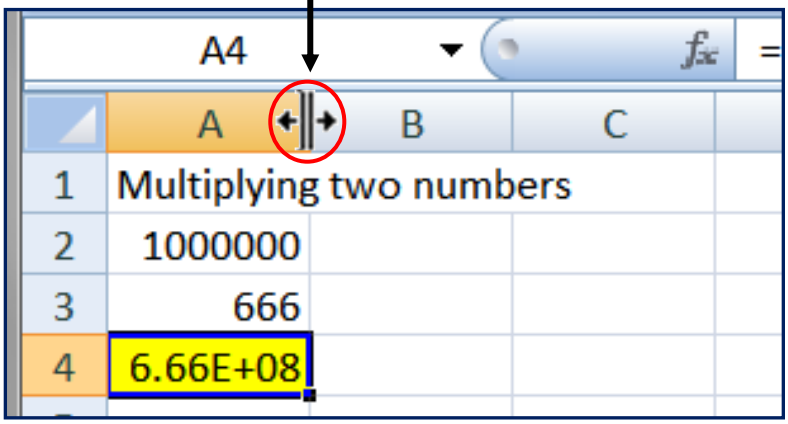
14.1b – Resizing Column Widths

IF you enter large numbers into cells A2 and A3 you may not get the result you were expecting. It may look like this:



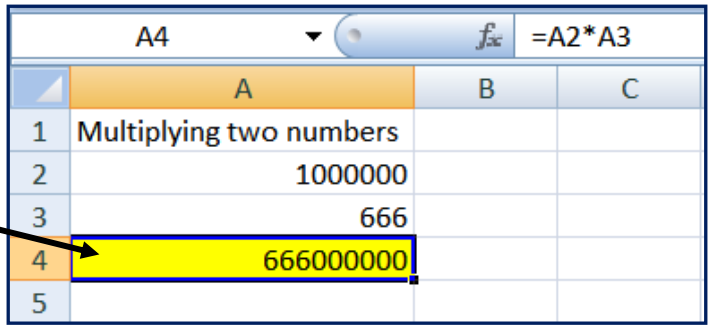
This tells you that the number is too big to fit into the column and you need to expand it.

Move the cursor to the end of column A like this.



Double click the left mouse button which will expand the column width to fit the contents of the longest item.

You should now be able to see all of the data.



Save the spreadsheet as **Data Analysis – Task A**

14.2 – Creating a simple Data Model

Task B

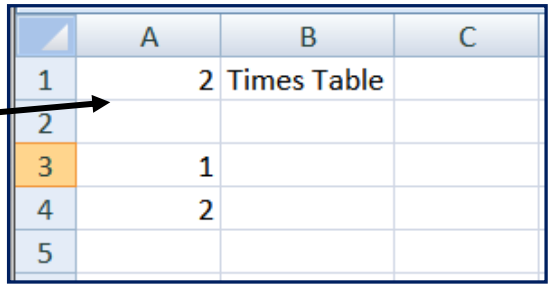
Create a spreadsheet to display the times table for any number you choose to enter.

Print your spreadsheet, showing values and formulae.

Save the spreadsheet as **Data Analysis – Task B**

How to do it:

For this task you need to copy out the spreadsheet shown in the picture here.



	A	B	C
1	2	Times Table	
2			
3	1		
4	2		
5			

You are going to create the **times table** in cells **A3** to **B12**.

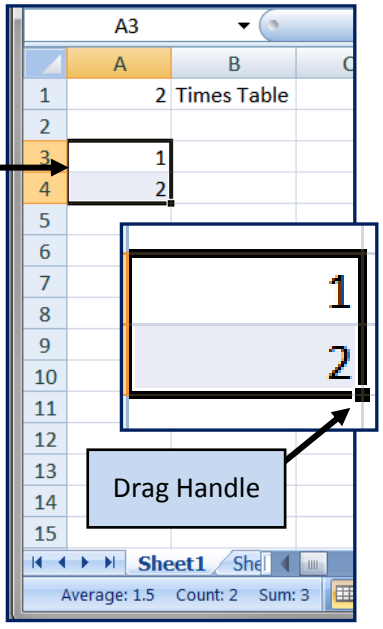
The cells in **column A** will hold the **number to multiply by** and those in **column B** will **hold the formulae** to calculate the answer.

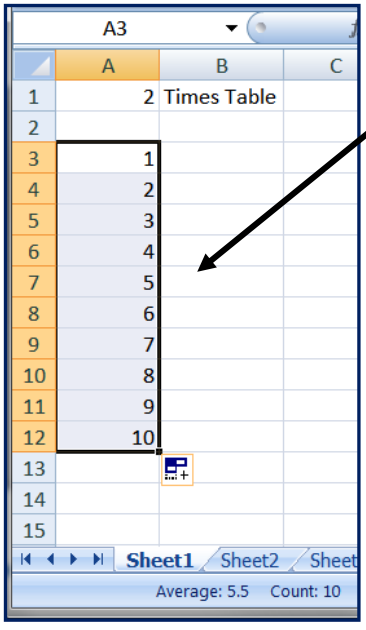
Autofil Function

Rather than manually filling in the numbers 3 to 10, **highlight the cells A3 and A4** as shown in the picture.

Find the **Drag Handle** in the bottom right corner of the two cells.

Click and hold the left mouse button on the drag handle and drag it down to cell **A12**.





This **replicates** (copies) the cell contents.

NOTE:

Excel is clever enough to realise that the numbers in cells A3 and A4 increase by 1 and uses this pattern as it copies the cells down to A12.

Adding the Formulae with an Absolute Cell Reference

Move the cursor into cell B3 and enter the following formulae:

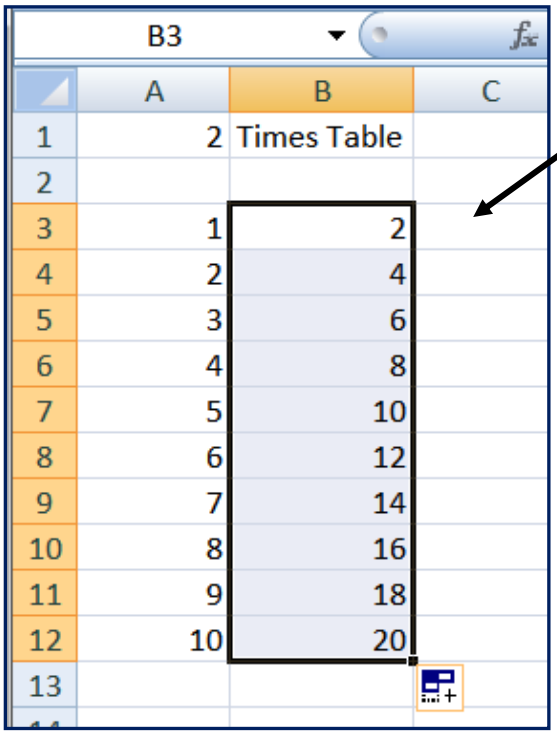
=A3*$A$1

NOTE:

The $ symbols used in the A1 cell reference tell excel to only use the contents of that particular cell when the formulae is replicated into cells B4 and B12.

This is known as 'Absolute Cell Referencing'.

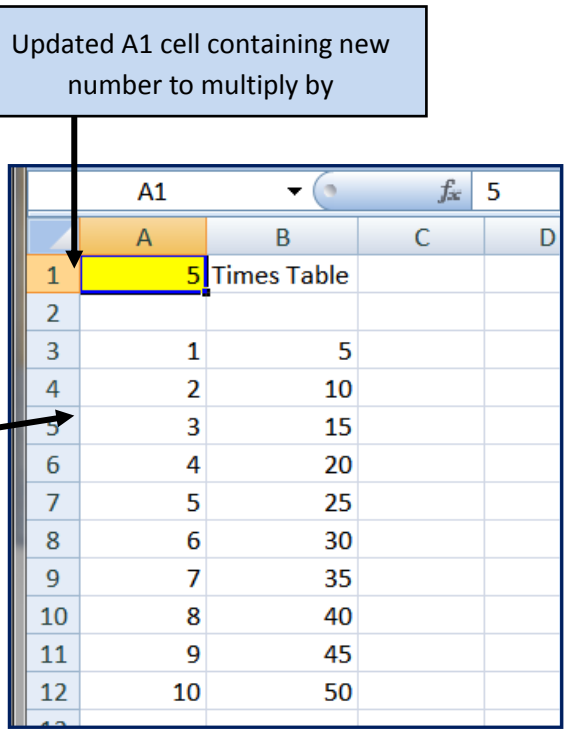
Use the Autofil Drag Handles in cell B3 to copy the formula into the cells down to B12.



The results should look like this:

You can see that this produces the correct results for the two times table.

To test the model, change the cell A1 to another number. Your model should adapt to this new number and re-calculate the updated times table answers.



Save the spreadsheet as **Data Analysis – Task B**

14.2b – Printing Spreadsheet Values

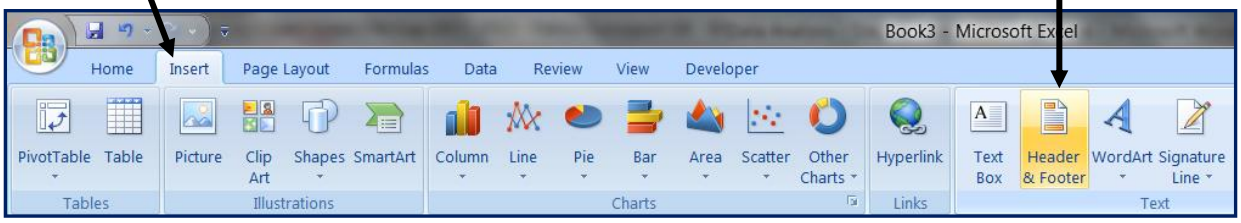
NOTE:

Whenever you are asked to print something (In any program – not just Excel) you MUST make sure you include in the Footer your:

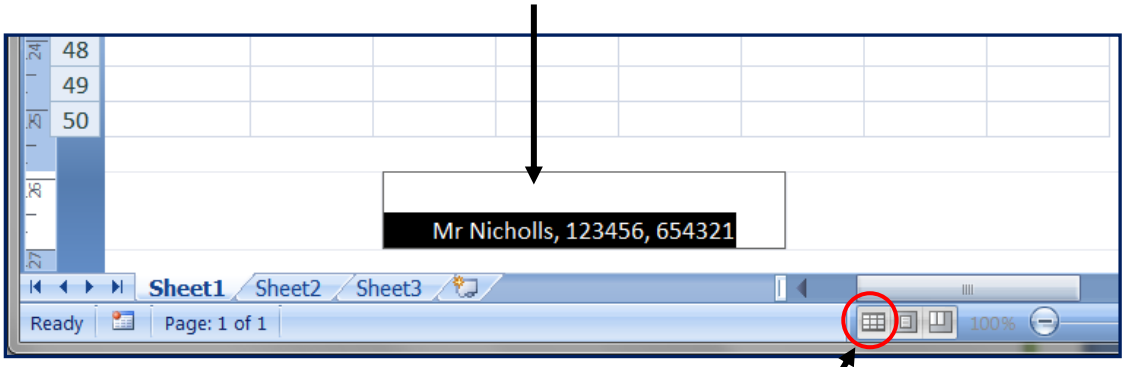
- **Full name**
- **Candidate number**
- **Centre Number.**

Adding name and detail to the footer

Click **Insert** then **Header and Footer** (This takes you to Page Layout view).



Scroll down into the **footer** and enter your details.

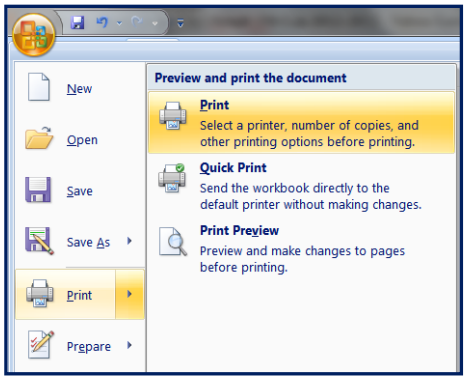


You can return to Normal View by clicking the **'Normal'** view option.

Printing Values View of the spreadsheet

Click the **Office Button**.

Click **Print** and the then **Print** again.

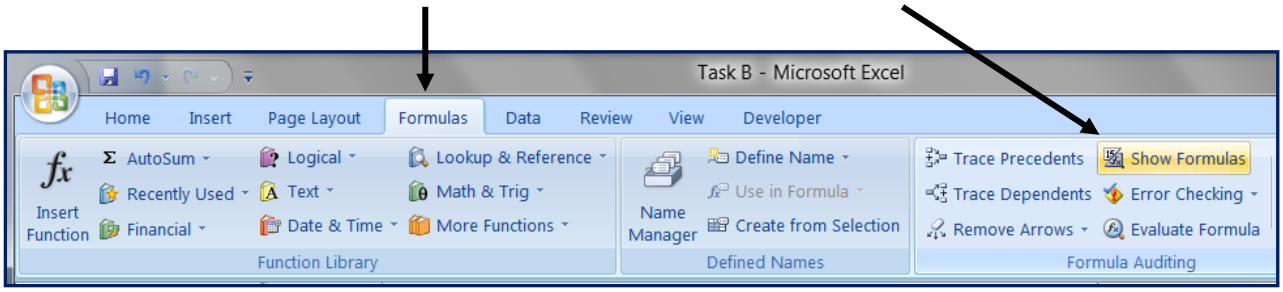


14.2c – Printing Spreadsheet Formulae

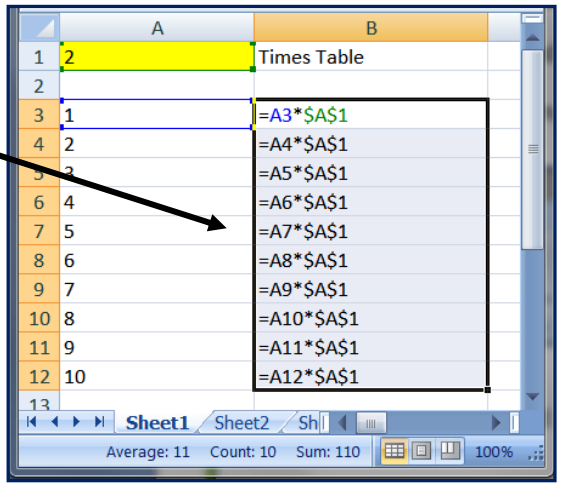
Viewing Spreadsheet Formulae

Sometimes you are asked to print off the **Formulae** used within a spreadsheet.

To do this, click the **Formulas tab** and find the **Show Formulas** icon.



Your spreadsheet should now display all of the formulae you have used.



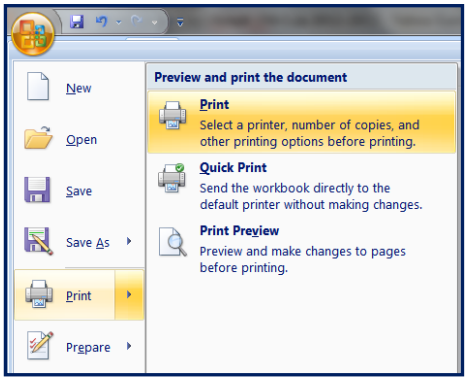
The formulae clearly show your use of **absolute and relative cell referencing** as well as the **calculations** used.

NOTE: Remember that absolute cell references are identified with the $ symbols.

Printing Formulae View of the spreadsheet

Click the **Office Button**.

Click **Print** and the then **Print** again.



NOTE:
To return to the view of the spreadsheet that shows the values, click on the Show Formulas icon again.

14.3– Accuracy of the data entry

When you are creating the spreadsheet within your exam it is **VITAL** that you **copy the model** in the question paper **EXACTLY** as shown.

DO NOT make any **improvements** or **enhancements** (such as colour formatting, alignments or extra columns/rows) unless asked to do so.

Likewise, make sure that any **data entered** is **EXACTLY** the same as the source spreadsheet or question paper. This applies to numbers, decimal places and formulas.

A large number of marks are lost due to careless data entry and rushing through the questions.

14.4 – Using Formulae

Simple mathematical operators can be used to:

- Add **(+)**– Also known as SUM
- Subtract **(-)**
- Divide **(/)**
- Multiply **(*)**
- Calculate Indices (powers) of a number **(^)**

Task C

Open the file **OPERATORS.CSV**.

Place **two numbers** of your choice in cells **B1** and **B2**. Calculate in cell:

- **B4**, the **sum** of the two numbers
- **B5**, the **difference** between the two numbers
- **B6**, the **product** of the two numbers
- **B7**, the contents of **cell B1divided** by the contents of **cell B2**
- **B8**, the contents of **cell B1** to the **power of** the contents of **cell B2**.

Check the calculations accuracy by comparing both the values and formulas to the images supplied. Call me over so that I can check the accuracy as well.

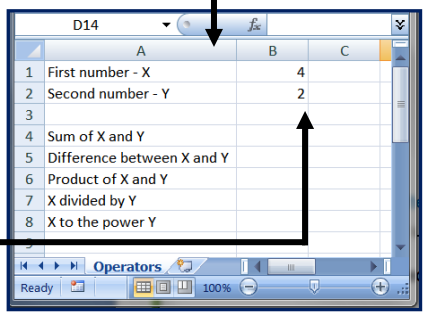
Save the spreadsheet as **Data Analysis – Task C**. (Make sure it is saved as an Excel Workbook)

How to do it:

Open **OPERATORS.CSV** from the Section 14 Files folder.

Extend the width of column A so that the **labels are fully visible** (See section 14.1b).

Move the cursor into cell **B1** and enter the number **4**, then into cell **B2** and enter the number 2.



NOTE: These numbers have been chosen so that you can easily check your calculation for accuracy.

1.AdditionCalculation:

Move the cursor into cell **B4**. You need to insert formula that **adds**(sums) the **contents of cell B1 and cell B2**.

Enter the following formula: **=B1+B2**

Hit the **<Enter>** key to set the formula.

2.Subtraction Calculation:

Move the cursor into cell **B5**. You need to insert formula that calculates the **difference between the two numbers**. (Difference being number 1 minus number 2)

Enter the following formula: **=B1-B2**

Hit the **<Enter>** key to set the formula.

3.Multiplication Calculation:

Move the cursor into cell **B6**. You need to insert formula that calculates the product **of the two numbers**. (Product means to multiply number 1 and number 2)

Enter the following formula: **=B1*B2**

Hit the **<Enter>** key to set the formula.

4. Division Calculation:

Move the cursor into cell **B7**. You need to insert formula that calculates the contents of cell **B1 divided by cell B2**.

Enter the following formula: **=B1/B2**

Hit the **<Enter>** key to set the formula.

5. Indices Calculation:

Move the cursor into cell **B8**. You need to insert formula that calculates the **contents of cell B1 to the power of the contents of cell B2**.

Enter the following formula: **=B1^B2**

Hit the **<Enter>** key to set the formula.

NOTE: The ^ symbol is usually found by pressing <Shift> and '6'.

Checking the spreadsheet for accuracy

Compare your spreadsheet to the image on the right to ensure that your calculations have worked.

Column A should be **resized** to display all data.

Column B should display the correct outcomes of each calculation.

	A	B
1	First number - X	4
2	Second number - Y	2
3		
4	Sum of X and Y	6
5	Difference between X and Y	2
6	Product of X and Y	8
7	X divided by Y	2
8	X to the power Y	16
9		

Select the Formulas tab, and then click on the Show formulas icon to display all of your formulae.

Again, these should be compared to those in the picture on the right:

	A	B
1	First number - X	4
2	Second number - Y	2
3		
4	Sum of X and Y	=B1+B2
5	Difference between X and Y	=B1-B2
6	Product of X and Y	=B1*B2
7	X divided by Y	=B1/B2
8	X to the power Y	=B1^B2
9		

Save spreadsheet as **Data Analysis–Task C.**(Make sure it is saved as Excel Workbook)

14.5 – Named Cells and Ranges

When an individual cell (or group of cells) is going to be used a number of times within the formulae of a spreadsheet, it is a good idea to give it a name.

Names given should be short and meaningful.

For example: If cell A5 contains the age of a person you can change the cell reference from A5 to actually naming it Age.

This is often easier to remember when using cell references as part of a formula.

Task D

Open the file **SALES.CSV**. This spreadsheet will be used to calculate a bonus payment to sales staff for a small company.

Name cell **B1 'Unit'**. Name cells **A5 to C7 'Rate'**. Names cells **B11 to G18 'Sold'**.

Save the spreadsheet as **Data Analysis – Task D**

How to do it:

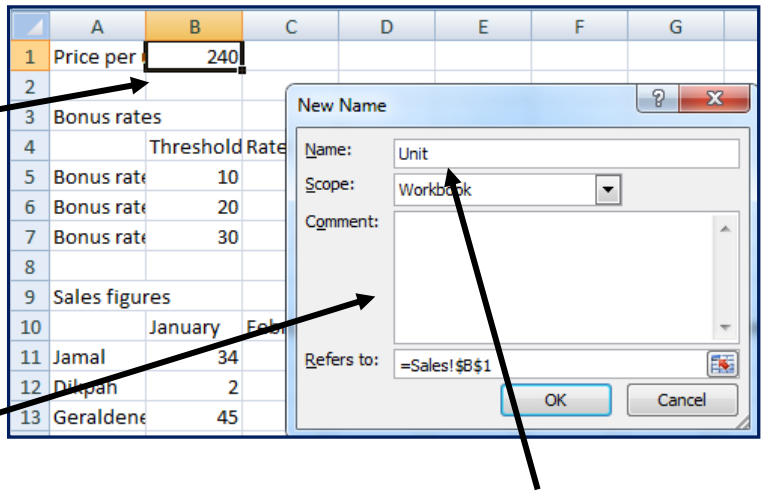
Naming Single Cells

Open the file **SALES.CSV** and find cell **B1**.

Right click the mouse on the cell to get a drop-down menu.

Select the option **Name a Range…**

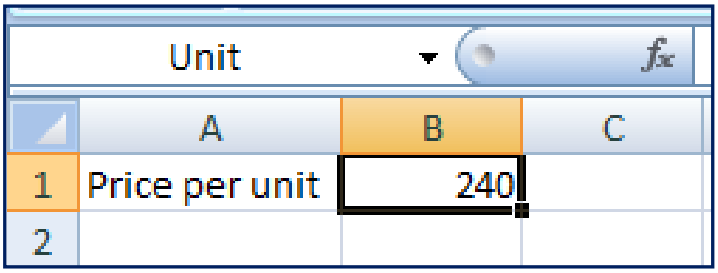
This will open the **New Name** window.



In the **Name:** box, overwrite the name the Excel suggests with the word **Unit**.

Click **OK**.

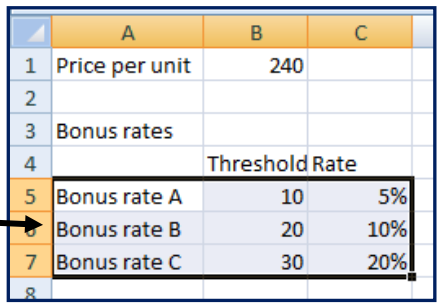
When you click on cell **B1**, you will now see that it has been renamed to **Unit**.



Naming Cell Ranges

To name the range of cells **(A5 to C7)** you must first highlight them.

To do this, click your mouse in cell **A5** then **hold down the mouse button and drag to cell C7**.

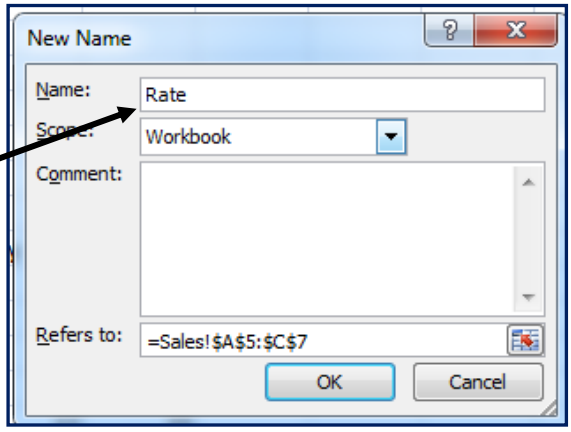


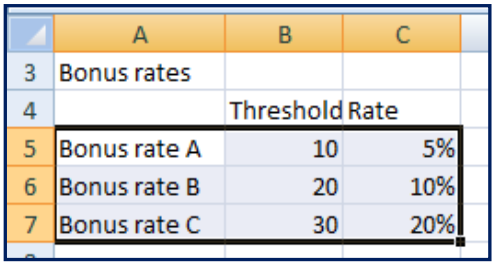
	A	B	C
1	Price per unit	240	
2			
3	Bonus rates		
4		Threshold	Rate
5	Bonus rate A	10	5%
6	Bonus rate B	20	10%
7	Bonus rate C	30	20%
8			

Right Click the mouse within the highlighted range to get the drop-down menu.

Change the contents of the **Name:** box to **Rate**.

Check that your New Name window looks like this screenshot before clicking on **OK**.



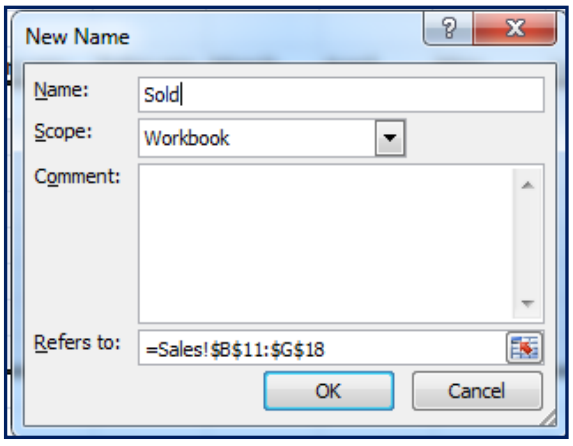


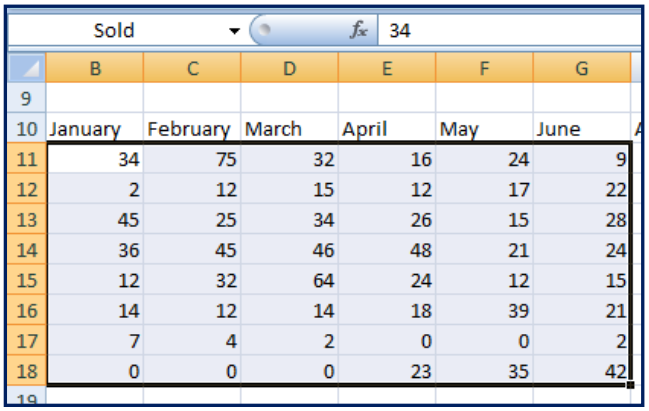
	A	B	C
3	Bonus rates		
4		Threshold	Rate
5	Bonus rate A	10	5%
6	Bonus rate B	20	10%
7	Bonus rate C	30	20%

Now when you highlight cells **A5 to C7**, you will see that they have collectively been renamed to **Rate**.

The final range can be created in the same way as above.

Highlight the cells **B11 to G18** and then rename them to **Sold**.





Sold | 34

	B	C	D	E	F	G
9						
10	January	February	March	April	May	June
11	34	75	32	16	24	9
12	2	12	15	12	17	22
13	45	25	34	26	15	28
14	36	45	46	48	21	24
15	12	32	64	24	12	15
16	14	12	14	18	39	21
17	7	4	2	0	0	2
18	0	0	0	23	35	42
19						

Save spreadsheet as **Data Analysis–Task D.**
(Make sure it is saved as Excel Workbook)

14.6 – Using Functions

All formulas in Excel starts with an = sign. However they vary in complexity. For example:

- Formulas can be **simple** and only use **mathematical operators** like **=A1 + A2**.
- Formulas can be **complex** and use **nested statements** (cover later in the book.)
- Formulas can include **functions** such as **SUM** or **AVERAGE**. Functions help us to more easily perform a particular calculation.

Task E

Copy this spreadsheet model and then calculate:

- The **total (SUM)** number of hours worked by all of these five people
- The **average** number of hours worked per person
- The **maximum** number of hours worked by any of these five people
- The **minimum** number of hours worked by any of these five people.

Save the spreadsheet as **Data Analysis – Task E**

	A	B
1	**Rate of Pay**	$12.80
2		
3	Name	Hours
4	Aaron Kane	26
5	Jeff Leathley	20
6	Jonathan Harrington	17
7	James Mitchell	4
8	Sue Gray	13
9	**Total:**	
10	**Average:**	
11	**Maximum:**	
12	**Minimum:**	

How to do it:

SUM Function

The **SUM** function is used to **add more than two numbers** together.

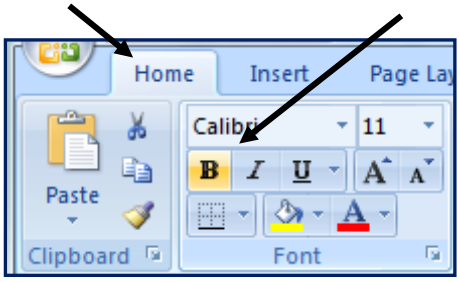
NOTE:

If you **only have two numbers to add** together (**A1 and A2** for example) then it is more efficient to simply add them **without** using the **SUM** function **(=A1 + A2)**.

However, if there were **three or more numbers to add**, using the **SUM** function is **quicker** and **more efficient**.

Open a new spreadsheet and copy the labels and values exactly as shown in the task.

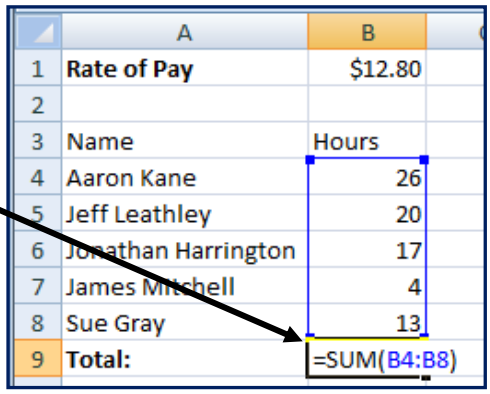
Select the **Home** tab and use the **Bold** icon to embolden the cells shown.



To find the total number of hours worked you need to click in cell **B9**.

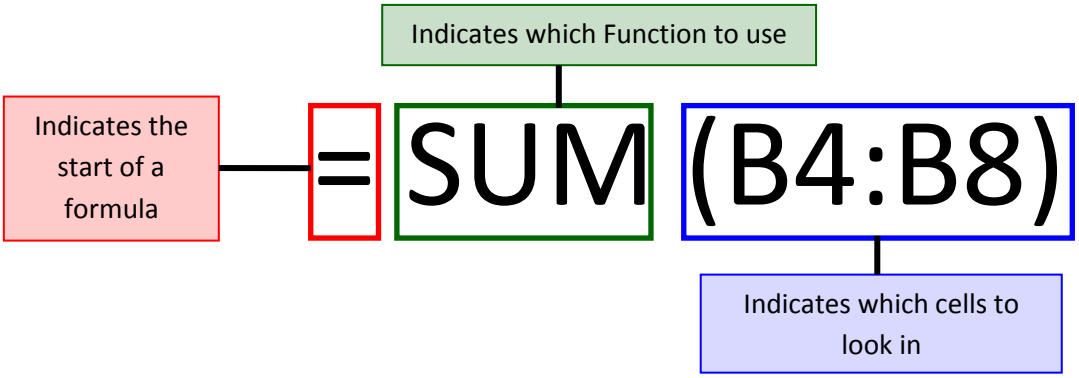
Enter the formula **=SUM(B4:B8)** the press **<Enter>.**

This should give the value **80**.



	A	B
1	**Rate of Pay**	$12.80
2		
3	Name	Hours
4	Aaron Kane	26
5	Jeff Leathley	20
6	Jonathan Harrington	17
7	James Mitchell	4
8	Sue Gray	13
9	**Total:**	=SUM(B4:B8)

Breakdown of the formula



Ways of using the SUM function

There are many ways of using the **SUM** function. Some of these ways are highlighted in the table below:

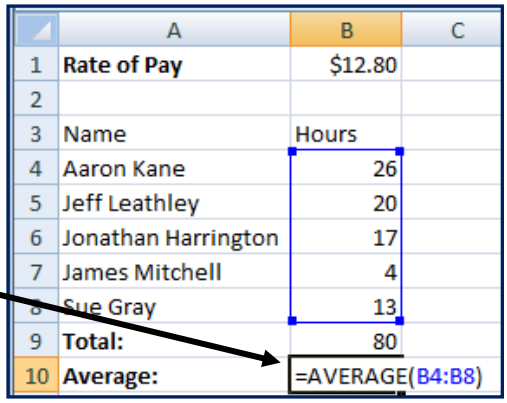
Function	Equivalent Formula	What it does
=SUM(B4:B8)	**=B4+B5+B6+B7+B8**	Adds up the contents the contents of all the cells in the range B4 to B8
=SUM(D3,D8,D12)	**=D3+D8+D12**	Adds up the contents of the cells D3, D8 and D12
=SUM(D5:D8,F2)	**=D5+D6+D7+D8+F2**	Adds up the contents of the cells in the range D5 to D8 and the contents of cell F2
=SUM(RangeName)	None	Adds up the contents of all the cells within a named range. This can be used with any named range

AVERAGE Function

To find the AVERAGE number of hours worked, click into cell **B10**.

Enter the formula **=AVERAGE(B4:B8)**.

This should give the value **16**.



Ways of using the AVERAGE function

There are many ways of using the **AVERAGE** function. Some of these ways are highlighted in the table below:

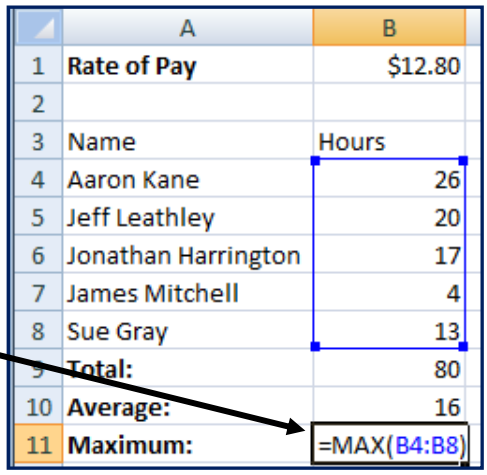
Function	Equivalent Formula	What it does
=AVERAGE(B4:B8)	**=(B4+B5+B6+B7+B8)/5**	Calculates the mean of all the cells in the range B4 to B8
=AVERAGE(D3,D8,D12)	**=(D3+D8+D12)/3**	Calculates the mean of the cells D3, D8 and D12
=AVERAGE(D5:D8,F2)	**=(D5+D6+D7+D8+F2)/5**	Calculates the mean of the cells in the range D5 to D8 and the contents of cell F2
=AVERAGE(RangeName)	None	Calculates the mean of all cells within a named range. This can be used with any named range

MAX Function

To find the person who worked the most **(MAX)** hours, click into cell **B11**.

Enter the formula **=MAX(B4:B8)**.

This should give the value **26**.



MIN Function

To find the person who worked the least **(MIN)** number of hours, click into cell **B12**.

Enter the formula **=MIN(B4:B8)**.

This should give the value **4**.

	A	B
1	**Rate of Pay**	$12.80
2		
3	Name	Hours
4	Aaron Kane	26
5	Jeff Leathley	20
6	Jonathan Harrington	17
7	James Mitchell	4
8	Sue Gray	13
9	**Total:**	80
10	**Average:**	16
11	**Maximum:**	26
12	**Minimum:**	=MIN(B4:B8)

Your final spreadsheet should look like this:

	A	B
1	**Rate of Pay**	$12.80
2		
3	Name	Hours
4	Aaron Kane	26
5	Jeff Leathley	20
6	Jonathan Harrington	17
7	James Mitchell	4
8	Sue Gray	13
9	**Total:**	80
10	**Average:**	16
11	**Maximum:**	26
12	**Minimum:**	4

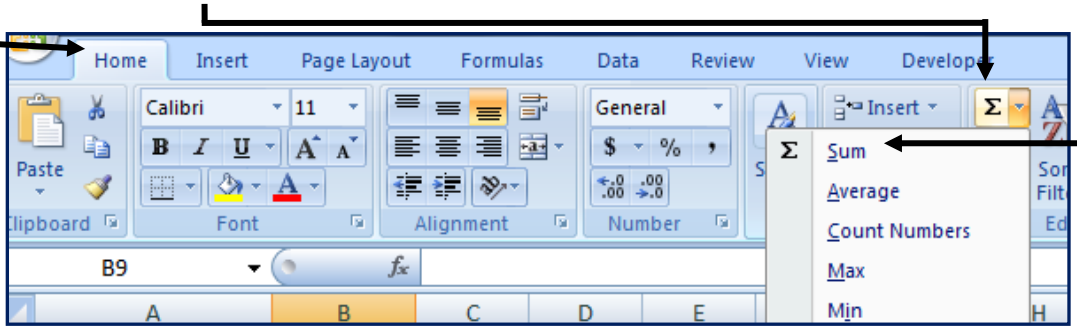
Save the spreadsheet as **Data Analysis – Task E** (Make sure it is saved as an Excel Workbook)

14.6a Alternative to typing in Functions

An alternative to typing in the functions manually is to use the **Functions Menu** found on the **Home Menu**.

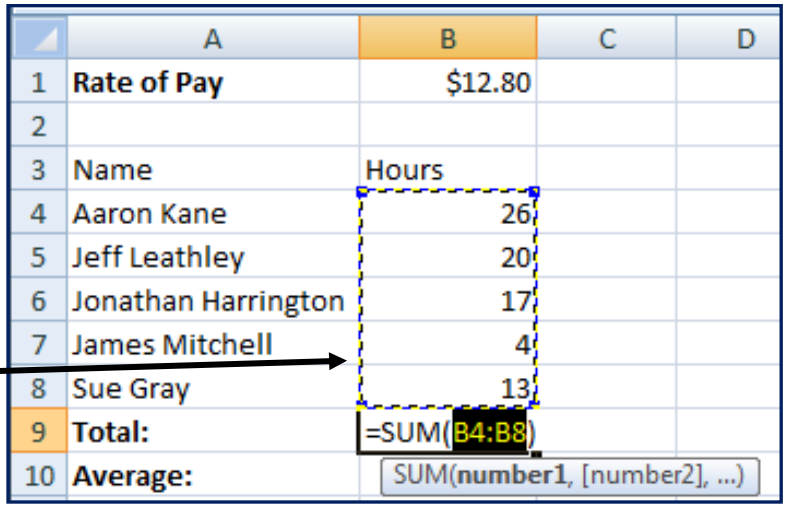
Open up the spreadsheet you saved in **Data Analysis – Task E** and delete cell **B9**.

Keep the cursor in cell **B9** and then select the **Home** tab. Click on the arrow to the right of the **Functions Icon**. Select the **AutoSum** icon.



This will place the **SUM** function into cell **B9** and attempt to work out which cells you wish to add up.

If it does not get the range correct you can highlight the cells to be added using your mouse.



Press **<Enter>** to accept the **AutoSum**.

NOTE: The Function Menu will also work with AVERAGE, MAX, MIN and most of the other functions that we discuss within this work booklet.

Activity 1

Open the file **TUCKSHOP.CSV**. Widen all columns that are too small to fit their contents.

In cells **B14 to B17**, calculate:

- The **total number of days** that all the students worked in the school shop.
- The **average number of days** worked.
- The **maximum and minimum** values.

Place your **name** in the document **footer**.

Print the spreadsheet showing the **values**.

Print the spreadsheet showing the **formulae used**.

Save the document as **Data Analysis – Activity 1**.

Task F

Aaron Kane did an extra four hours' work. Change the spreadsheet you saved in **Task E** to show the new figures.

NOTE: The manager wants to see the **average number of hours** worked displayed as:

- An **integer** value
- **Rounded** to the nearest **whole hour**.

Save the spreadsheet as **Data Analysis – Task F**

How to do it:

Open the file you saved as **Data Analysis – Task E**.

Change the contents of cell **B4** to **'30'** to add the four extra hours that he worked.

This gives an overall average value of **16.8 hours**.

	A	B
1	**Rate of Pay**	$12.80
2		
3	Name	Hours
4	Aaron Kane	30
5	Jeff Leathley	20
6	Jonathan Harrington	17
7	James Mitchell	4
8	Sue Gray	13
9	**Total:**	84
10	**Average:**	16.8
11	**Maximum:**	30
12	**Minimum:**	4

	A	B	C	D
1	**Rate of Pay**	$12.80		
2				
3	Name	Hours		
4	Aaron Kane	30		
5	Jeff Leathley	20		
6	Jonathan Harrington	17		
7	James Mitchell	4		
8	Sue Gray	13		
9	**Total:**	84	**Integer**	**Rounding**
10	**Average:**	16.8		
11	**Maximum:**	30		
12	**Minimum:**	4		

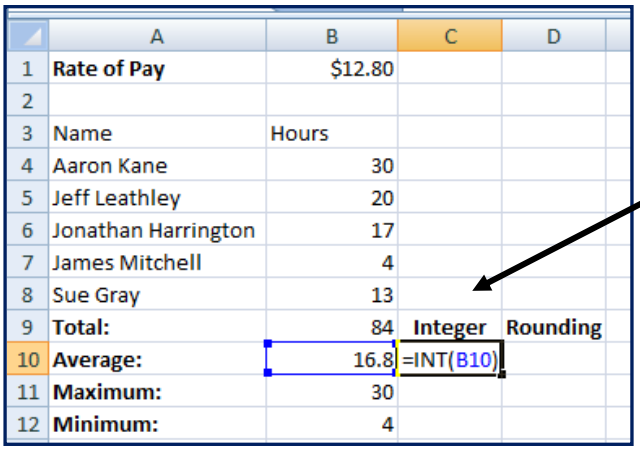
Move the cursor into cell **C9** and enter the text **'Integer'**. Move the cursor into cell **D9** and enter the text **'Rounding'**.

To get the first value requested by the manager **(Integer)**, we have to set cell **B10** to hold an **Integer** value.

INT Function

In mathematics, an **integer** is the word used to describe a **Whole Number (with no decimals or fractions)**.

NOTE: In Excel, the INT function takes a number and removes all digits after the decimal point.



	A	B	C	D
1	**Rate of Pay**	$12.80		
2				
3	Name	Hours		
4	Aaron Kane	30		
5	Jeff Leathley	20		
6	Jonathan Harrington	17		
7	James Mitchell	4		
8	Sue Gray	13		
9	**Total:**	84	**Integer**	**Rounding**
10	**Average:**	16.8	=INT(B10)	
11	**Maximum:**	30		
12	**Minimum:**	4		

Move the cursor into cell **C10** and enter the formula:

=INT(B10).

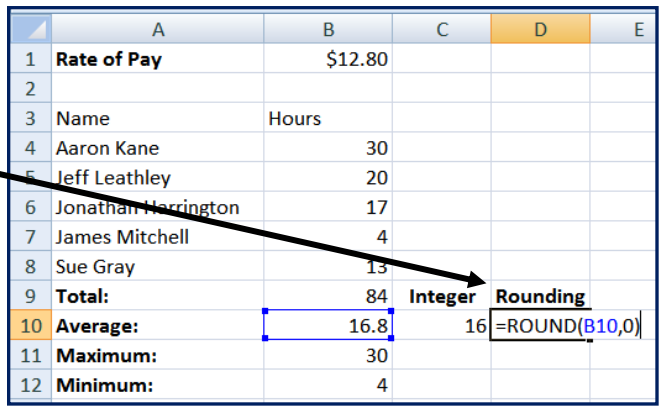
This should successfully remove the .8 and leave you with the whole number of 16.

ROUND Function

Move the cursor into cell **D10** and enter the following formula:

=ROUND(B10,0)

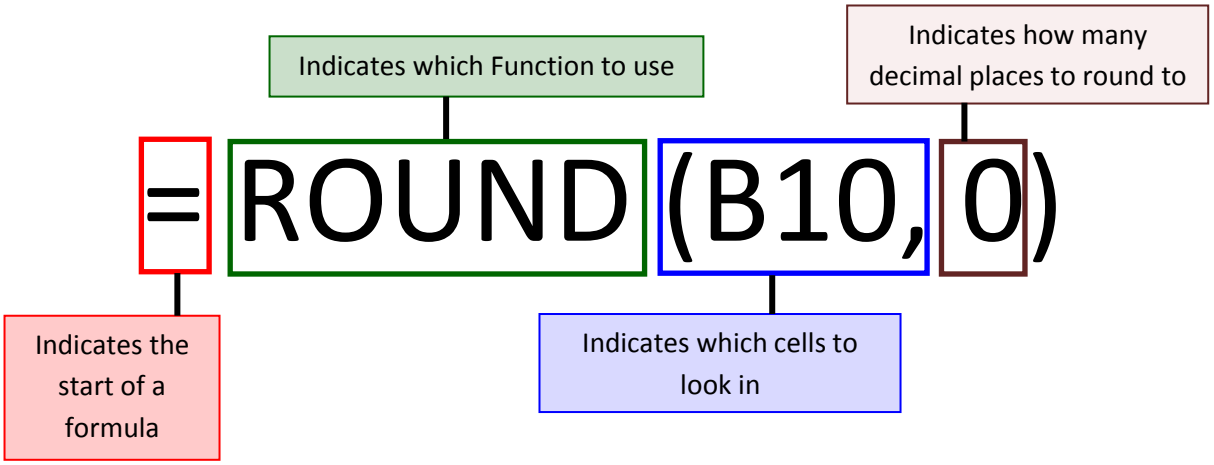
This uses the **ROUND** function which takes the contents of cell **B10** and **rounds** the number to **0 decimal places**.



	A	B	C	D	E
1	**Rate of Pay**	$12.80			
2					
3	Name	Hours			
4	Aaron Kane	30			
5	Jeff Leathley	20			
6	Jonathan Harrington	17			
7	James Mitchell	4			
8	Sue Gray	13			
9	**Total:**	84	**Integer**	**Rounding**	
10	**Average:**	16.8	16	=ROUND(B10,0)	
11	**Maximum:**	30			
12	**Minimum:**	4			

NOTE: IF the decimal point is 5 or above, the formula will round the digit to the nearest whole number. (16.8 would become 17 for example)

Breakdown of the formula



	A	B	C	D
1	**Rate of Pay**	$12.80		
2				
3	Name	Hours		
4	Aaron Kane	30		
5	Jeff Leathley	20		
6	Jonathan Harrington	17		
7	James Mitchell	4		
8	Sue Gray	13		
9	**Total:**	84	**Integer**	**Rounding**
10	**Average:**	16.8	16	17
11	**Maximum:**	30		
12	**Minimum:**	4		

The spreadsheet should look like this.

Ways of using the ROUND function

There are many ways of using the **ROUND** function. Some of these ways are shown in the table below.

NOTE: All ROUND functions are used in cell A1 which contain the number 62.5512.

Function	Result of Rounding	What it does
=ROUND(A1,2)	**62.55**	Rounds the contents of A1 to two decimal places.
=ROUND(A1,1)	**62.6**	Rounds the contents of A1 to one decimal place. The digit '5' has forced the previous figure to be rounded up.
=ROUND(A1,0)	**63**	Rounds the contents of A1 to zero decimal places. The digit '5' in the 62.5512 has forced the previous figure to be rounded up.
=ROUND(A1,-1)	**60**	Rounds the contents of A1 to the nearest 10.
=ROUND(A1,-2)	**100**	Rounds the contents of cell A1 to the nearest 100.

Save the spreadsheet as **Data Analysis – Task F**
(Make sure it is saved as an Excel Workbook)

Activity 2

Create a new spreadsheet model to calculate:

- The **whole number** of 375.56411
- 375.56411 rounded to **two decimal places**
- 375.56411 rounded to the **nearest whole number**
- 375.56411 rounded to the **nearest ten**
- 375.56411 rounded to the **nearest hundred**
- 375.56411 rounded to the **nearest thousand**

Save the document as **Data Analysis – Activity 2**.

Task G

Open the file **PROJECT.CSV**.

This file lists some workers and the number of jobs they have still to finish for a project.

Place a formula in cell **A22** to **count the number of workers** that still have **jobs to be finished** for the project.

Place a formula in cell **A24** to **count the number of workers** on the project.

Save the spreadsheet as **Data Analysis – Task G**

How to do it:

Open the **PROJECT.CSV** file.

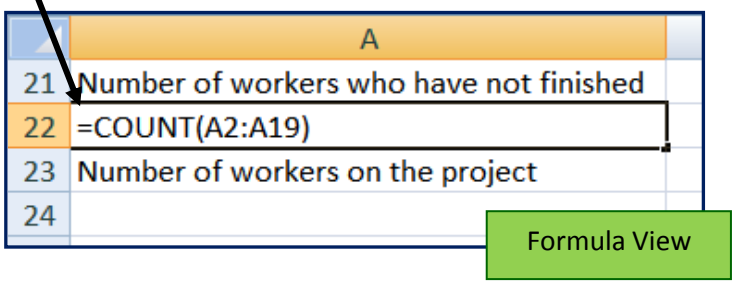
For this task you will need to use functions that **COUNT** different values.

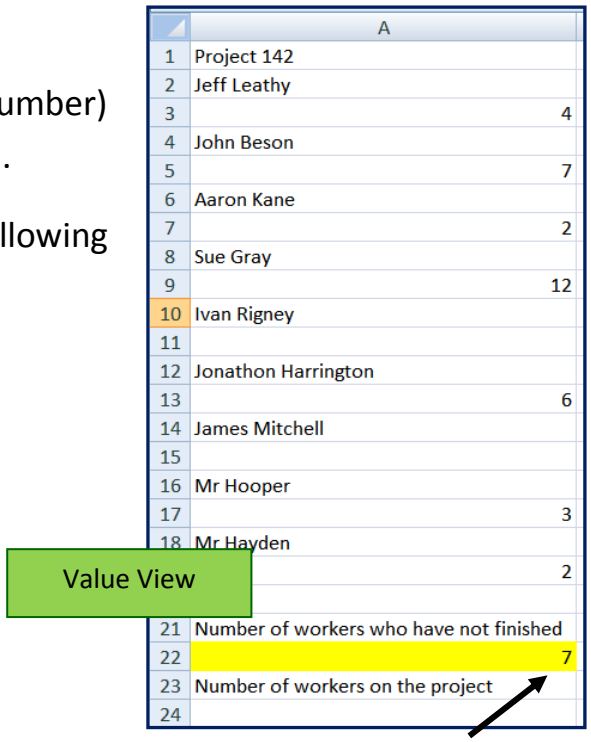
COUNT Function

It is possible to count the number of numeric (number) values in a spreadsheet using the **COUNT** function.

Place the cursor into cell **A22** and enter the following formula:

=COUNT(A2:A19).





This will look at the range **A2 to A19** and count the cells with numbers in them. The outcome of the **COUNT** should be **7**.

NOTE: Any cells that contain text or a combination of text and numbers will be ignored.

COUNTA Function

The **COUNTA** function works in a similar way to the **COUNT** function with a slight difference.

Rather than just being able to count the number of numeric values (like the COUNT function), **COUNTA** can count the number of numeric **OR** text vales in a cell.

It will **NOT count** any **empty** cells.

NOTE: In Excel there is not a function that can count text values while ignoring numeric values. Because of this our formula will have to include both the COUNTA and COUNT functions to calculate the number of workers on the project.

Place the cursor in cell **A24** and enter the following formula:

=COUNTA(A2:A19) –COUNT(A2:A19)

This will look at **A2 to A19** and **count** all the cells with **text OR numbers** in them. It will then **subtract** the **number of cells** with **ONLY numeric values** in them to leave only the cells with text in them.

It should give the value **9** and look like this.

	A
21	Number of workers who have not finished
22	=COUNT(A2:A19)
23	Number of workers on the project
24	=COUNTA(A2:A19)-COUNT(A2:A19)

Formula View

	A
1	Project 142
2	Jeff Leathy
3	4
4	John Benson
5	7
6	Aaron Kane
7	2
8	Sue Gray
9	12
10	Ivan Rigney
11	
12	Jonathon Harrington
13	6
14	James Mitchell
15	
16	Mr Hooper
17	3
18	Mr Hayden
19	2
20	
21	Number of workers who have not finished
22	7
23	Number of workers on the project
24	9

Value View

Save the spreadsheet as **Data Analysis – Task G**
(Make sure it is saved as an Excel Workbook)

Activity 3

Open the **CLASSLIST.CSV** file.

This spreadsheet lists all the students in a class. If a student has attended any clubs during the year, the number of times they have attended is recorded in the cell below their name.

Place a formula in cell **A71** to **count** the **number of students in the class**.

Place a formula in cell **A74** to count the **number of students who have attended extra clubs this year**.

Save the document as **Data Analysis – Activity 3**.

COUNTIF Function

Task H

Open the file **STAFF.CSV**.

This file lists some workers on another project and lists each workers job.

Name the cells **B3 to B21 'Job'**

Place a formula in cells **B24 to B28** to **count how many of each type of worker are employed** on the project.

Place a formula that uses **absolute cell referencing** in cell **B31** to **count** the **number of employees** with **less than five years experience**.

Place a formula that uses **absolute cell referencing** in cell **B32** to **count** the **number of employees** with **ten years or more experience**.

Save the spreadsheet as **Data Analysis – Task H**

How to do it:

Open the **STAFF.CSV** file.

For this task you will need to count how many people have each different type of job.

Name cells **B3 to B21** to **'Job'**.

Place the cursor in cell **B24**.

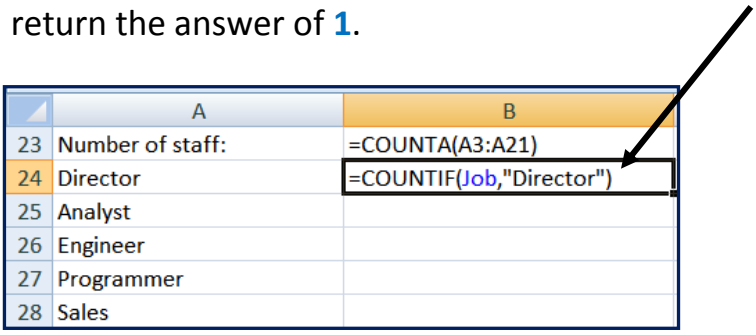
The function needed for this task is **COUNTIF**. This looks at the cells within a given range and counts the number of cells in that range that meet a certain condition.

NOTE: The condition can be a number, text, an inequality or a cell reference.

There are a number of ways the **COUNTIF** function can be used. Any of the formula in the table below will work:

Function	What it does
=COUNTIF($B$3:$B$21, "Director")	Counts the number of cells in the range B3 to B21 that contain the word 'Director'.
=COUNTIF(Job, "Director")	Counts the number of cells in the named range 'Job' (B3 to B21) that contain the word 'Director'.
=COUNTIF($B$3:$B$21, A24)	Counts the number of cells in the range B3 to B21 that contain the same text as the contents of cell A24.
=COUNTIF(Job, A24	Counts the number of cells in the named range 'Job' (B3 to B21) that contain the same text as the contents of cell A24.

Choose any of the above formula and make sure that it works. The formula should return the answer of **1**.



	A	B
23	Number of staff:	=COUNTA(A3:A21)
24	Director	=COUNTIF(Job,"Director")
25	Analyst	
26	Engineer	
27	Programmer	
28	Sales	

Repeat the formula to count how many workers carry out the other type of jobs. Your outcome should be the same as the screenshot below.

Formula View

		B
23	Number of staff:	=COUNTA(A3:A21)
24	Director	=COUNTIF(Job,"Director")
25	Analyst	=COUNTIF(Job,"Analyst")
26	Engineer	=COUNTIF(Job,"Engineer")
27	Programmer	=COUNTIF(Job,"Programmer")
28	Sales	=COUNTIF(Job,"Sales")

Value View

23	Number of staff:	19
24	Director	1
25	Analyst	2
26	Engineer	4
27	Programmer	8
28	Sales	4

NOTE: A quick count of the numbers should add up to a total of 19 (Number of staff)

To count the number of employees with **less than five years experience**, place the cursor into cell **B31**.

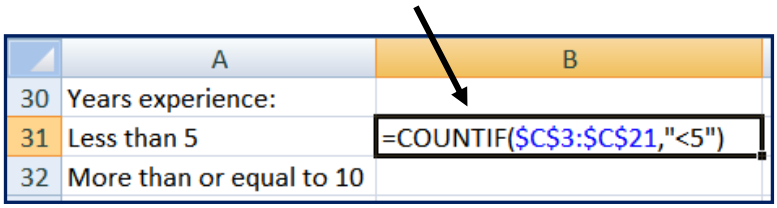
Enter the following formula:

=COUNTIF($C$3:$C$21, "<5")

This will look at the range **C3 to C21** and count the cells with a **number of less than five**.

NOTE: The speech marks " " around the <5 are needed to tell Excel that it is dealing with another formula. If you fail to type in the speech marks, Excel will look for the symbols <5.

The spreadsheet should return the value of **7**.



	A	B
30	Years experience:	
31	Less than 5	=COUNTIF($C$3:$C$21,"<5")
32	More than or equal to 10	

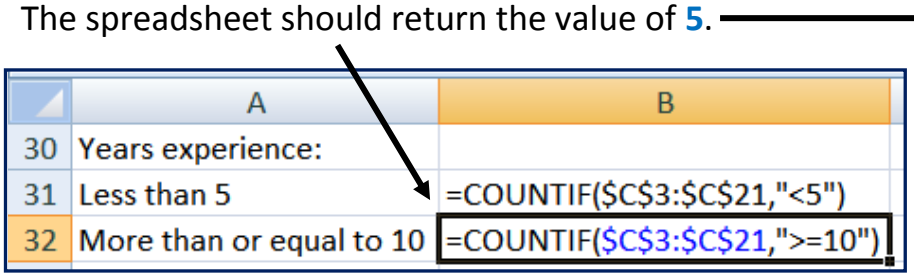
To count the number of employees with **experience of ten years or more**, place the cursor in cell **B32**.

Enter the following formula:

=COUNTIF($C$3:$C$21, ">=10")

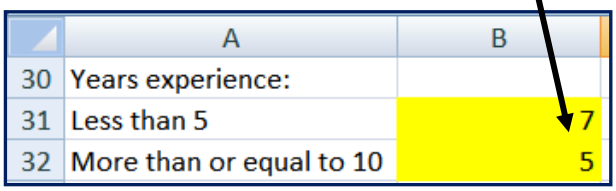
This will look at the range **C3 to C21** and count cells with a **number of ten or greater**.

The spreadsheet should return the value of **5**.



	A	B
30	Years experience:	
31	Less than 5	=COUNTIF($C$3:$C$21,"<5")
32	More than or equal to 10	=COUNTIF($C$3:$C$21,">=10")

Save the spreadsheet as **Data Analysis – Task H**
(Make sure it is saved as an Excel Workbook)



	A	B
30	Years experience:	
31	Less than 5	7
32	More than or equal to 10	5

Activity 4

Open the file that you saved in **Activity 3**.

This spreadsheet lists all the students in a class. Next to each student's name is the colour of the house that they are in.

Place a formula in cells E2 to E5 that use both **absolute** and **relative cell referencing** and **count** the **number of students in each house**.

Place a formula in cell **E7** to count the **number of students with less than five clubs**.

Place a formula in cell **E87** to count the **number of students with 12 or more clubs**.

Save the document as **Data Analysis – Activity 4**.

NOTE:

NOW YOU SHOULD COMPLETE THE EXTRA COUNTIF FUNCTION TASKS AS PRACTICE

SUMIF Function

SUMIF works in a similar way to **COUNTIF**.

SUMIF compares each value in a range of cells and if the value matches the given condition it will add another related cell to give a running total.

Task I

Open the file that you saved in **Data Analysis – Task H**.

Insert the following labels:

- **'Total experience for:'** into cell **A34**
- **'Programmer'** into cell **A35**
- **'Engineer'** into cell **A36**.

Place a formula into cell **B35** that uses both **absolute** and **relative cell referencing** to calculate the **number of years' experience** for the **programmers**.

Place a formula into cell **B36** that uses both **absolute** and **relative cell referencing** to calculate the **number of years' experience** for the **engineers**.

Save the spreadsheet as **Data Analysis – Task I**

How to do it:

Add the labels into cells **A34**, **A35** and **A36** as required by the task.

	A
33	
34	Total experience for:
35	Programmer
36	Engineer

Move the cursor into cell **B35** and enter the following formula:

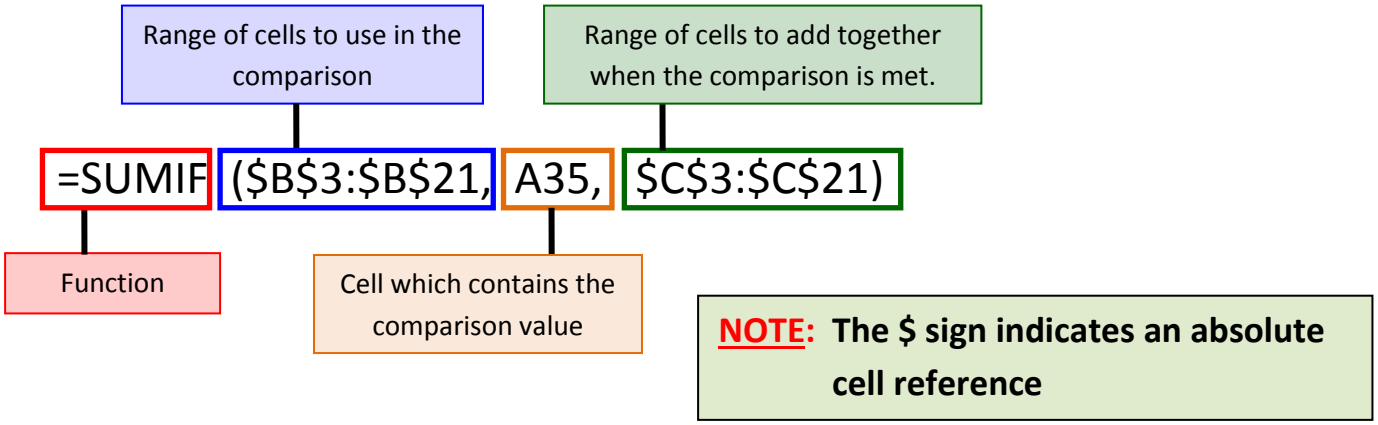
=SUMIF($B$3:$B$21,A35,$C$3:$C$21)

This will look at the **contents of each cell** in the **range B3 to B21** and then **compare their values** to the **contents of cell A35** (A35 contains the text 'Programmer').

When the **contents** of the **range B3 to B21** **match** the **contents of cell A35**, the **corresponding values of cells C3 to C21** will be **added** together.

	A	B
34	Total experience for:	
35	Programmer	=SUMIF($B$3:$B$21,A35,$C$3:$C$21)
36	Engineer	
37		

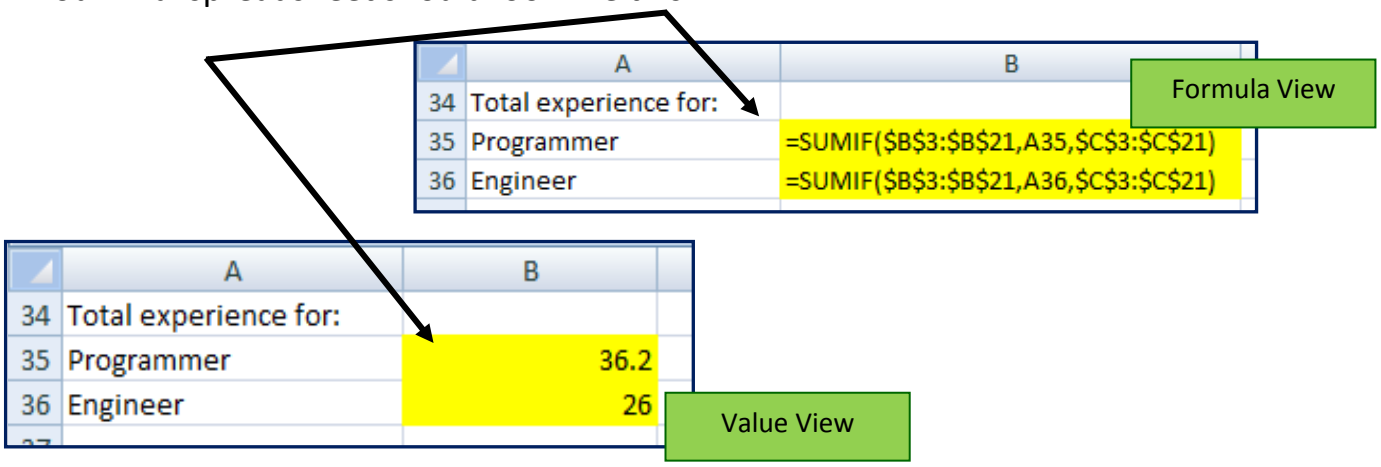
Breakdown of the formula



To total the years' experience for the engineers, place the cursor into cell **B36** and enter the following formula:

=SUMIF($B$3:$B$21,A36,$C$3:$C$21)

Your final spreadsheet should look like this:



Save the spreadsheet as **Data Analysis – Task I**
(Make sure it is saved as an Excel Workbook)

Activity 5

Open the file **CLUBS.CSV**.

Insert a formula into cell **B37** that uses both **absolute** and **relative referencing** to calculate the **number of clubs attended** by students in **red house**.

Replicate this formula into cells **B38 to B40** for the other coloured houses.

Save the document as **Data Analysis – Activity 5**.

NOTE:

NOW YOU SHOULD COMPLETE THE EXTRA SUMIF FUNCTION TASKS AS PRACTICE

SUMIF Function with NOT criteria

NOT criteria allow you to **exclude data** from your calculations.

NOT criteria are entered as this symbol **<>**

For example, you might to total shot put distances that were not recorded as no throws.

Task J

Open the file **Games.CSV**.

In cell **K8** enter a formula that will **add together the distances of all 6 throws** for that athlete.

Make sure that the function **does not include any distances that have been recorded as NT** (no throw)

Use **autofil** to replicate the formulae down to cell **K80**

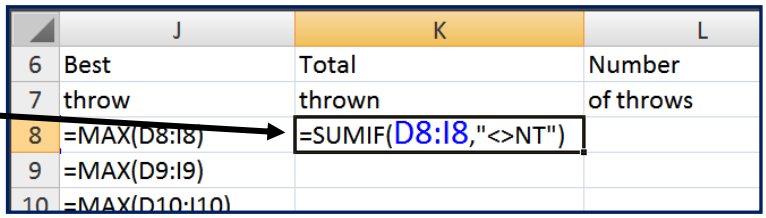
Save the spreadsheet as **Data Analysis – Task J**

How to do it:

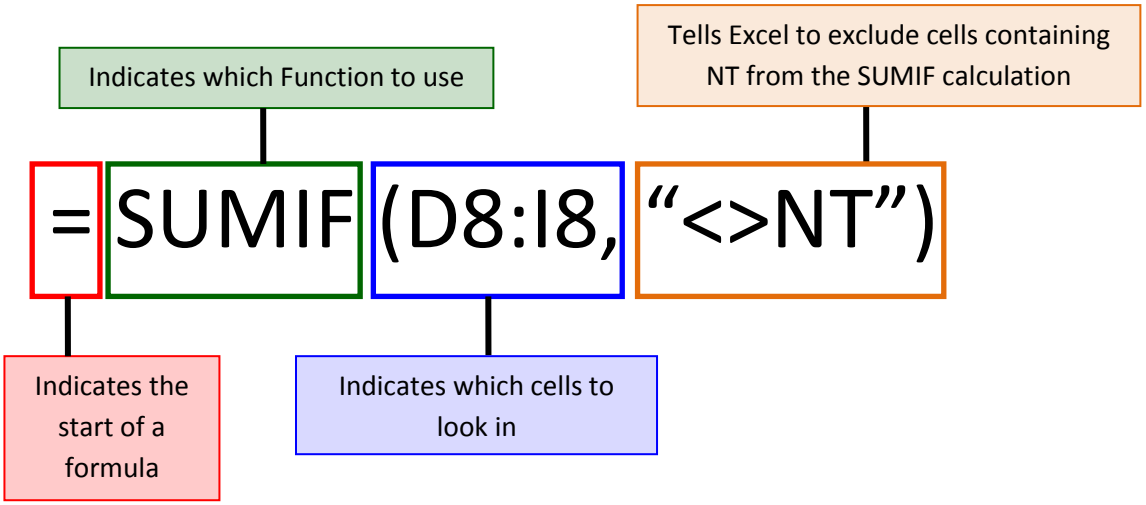
Open the file **Games.csv**.

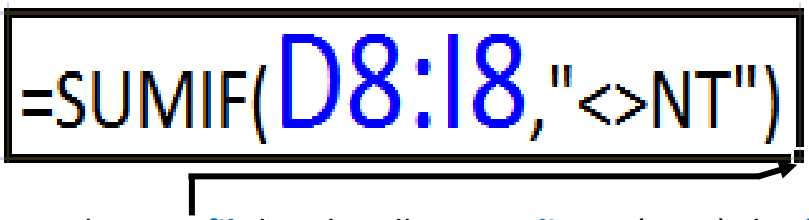
Place the cursor into cell **K8** and enter the following formulae:

=SUMIF(D8:I8,"<>NT")



Breakdown of the formula





Use the **autofil** drag handle to **replicate** (copy) the **formula** all the way down to cell **K80**.

Formula View

	K
71	=SUMIF(D71:I71,"<>NT")
72	=SUMIF(D72:I72,"<>NT")
73	=SUMIF(D73:I73,"<>NT")
74	=SUMIF(D74:I74,"<>NT")
75	=SUMIF(D75:I75,"<>NT")
76	=SUMIF(D76:I76,"<>NT")
77	=SUMIF(D77:I77,"<>NT")
78	=SUMIF(D78:I78,"<>NT")
79	=SUMIF(D79:I79,"<>NT")
80	=SUMIF(D80:I80,"<>NT")

Value View

	K
71	117.91
72	69.34
73	117.56
74	103.88
75	102.8
76	103.5
77	110.68
78	67.71
79	87.02
80	99.72

Save the spreadsheet as **Data Analysis – Task J**
(Make sure it is saved as an Excel Workbook)

COUNTIF Function with NOT criteria

NOT criteria works in exactly the same way with a **COUNTIF** function.

Remember that NOT criteria are entered as this symbol **<>**

An example would be to count the number of shot put throws that were not recorded as no throws.

Task K

Open the file that you saved in **Data Analysis – Task J**.

In cell **L8** enter a formula that will **count the number of throws** for that athlete that **were not recorded as NT** (no throw).

Use **autofil** to replicate the formulae down to cell **L80**

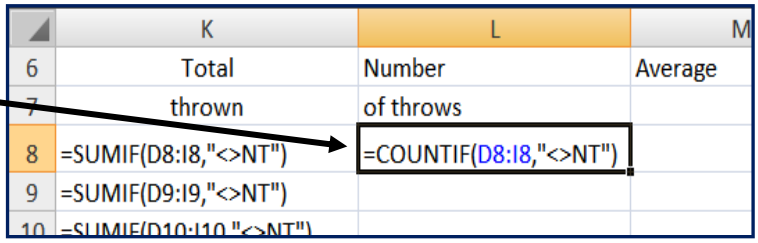
Save the spreadsheet as **Data Analysis – Task K**

How to do it:

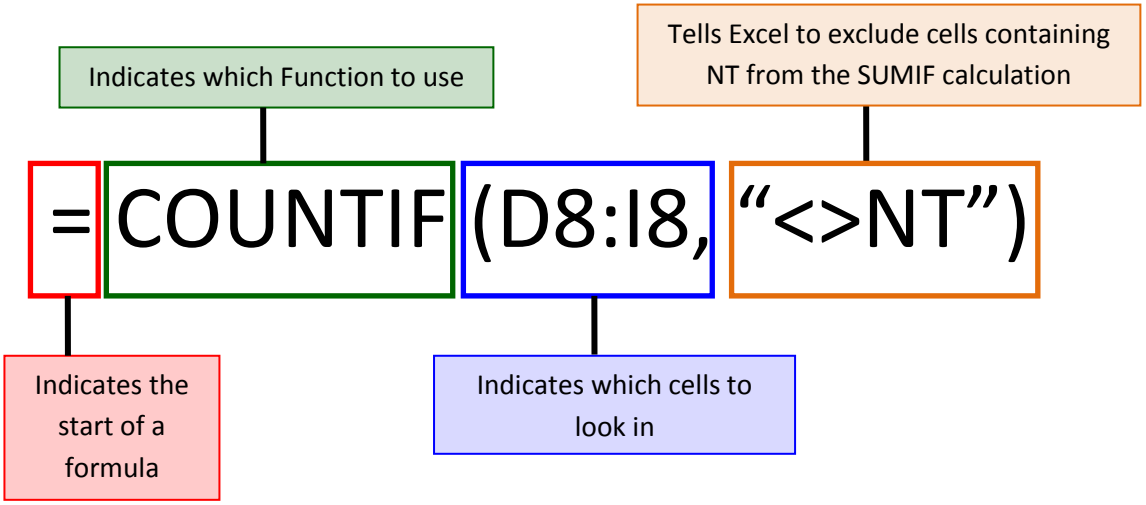
Open the file **Data Analysis – Task J**

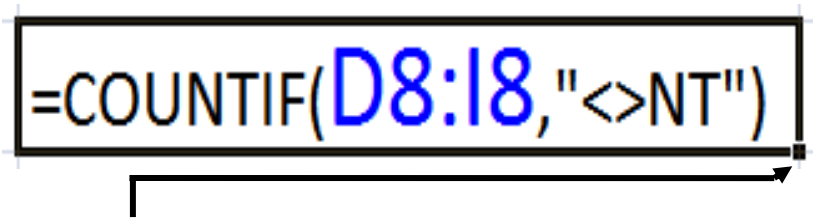
Place the cursor into cell **L8** and enter the following formulae:

=COUNTIF(D8:I8,"<>NT")



Breakdown of the formula





Use the **autofil** drag handle to **replicate** (copy) the **formula** all the way down to cell **K80**.

Formula View

	L
71	=COUNTIF(D71:I71,"<>NT")
72	=COUNTIF(D72:I72,"<>NT")
73	=COUNTIF(D73:I73,"<>NT")
74	=COUNTIF(D74:I74,"<>NT")
75	=COUNTIF(D75:I75,"<>NT")
76	=COUNTIF(D76:I76,"<>NT")
77	=COUNTIF(D77:I77,"<>NT")
78	=COUNTIF(D78:I78,"<>NT")
79	=COUNTIF(D79:I79,"<>NT")
80	=COUNTIF(D80:I80,"<>NT")

Value View

	L
71	6
72	4
73	6
74	5
75	6
76	6
77	6
78	4
79	5
80	5

Save the spreadsheet as **Data Analysis – Task K**
(Make sure it is saved as an Excel Workbook)

Activity 6

Open the file that you saved in **Data Analysis – Task J**.

Place a formula into cell **M8** to find the **average distance of the athlete's 6 throws**.

Use **NOT criteria** to make sure that the average distances **do not include any no throws** (NT).

Use **autofil** to replicate the formulae down to cell **M80**.

Save the document as **Data Analysis – Activity 6**.

IF Function

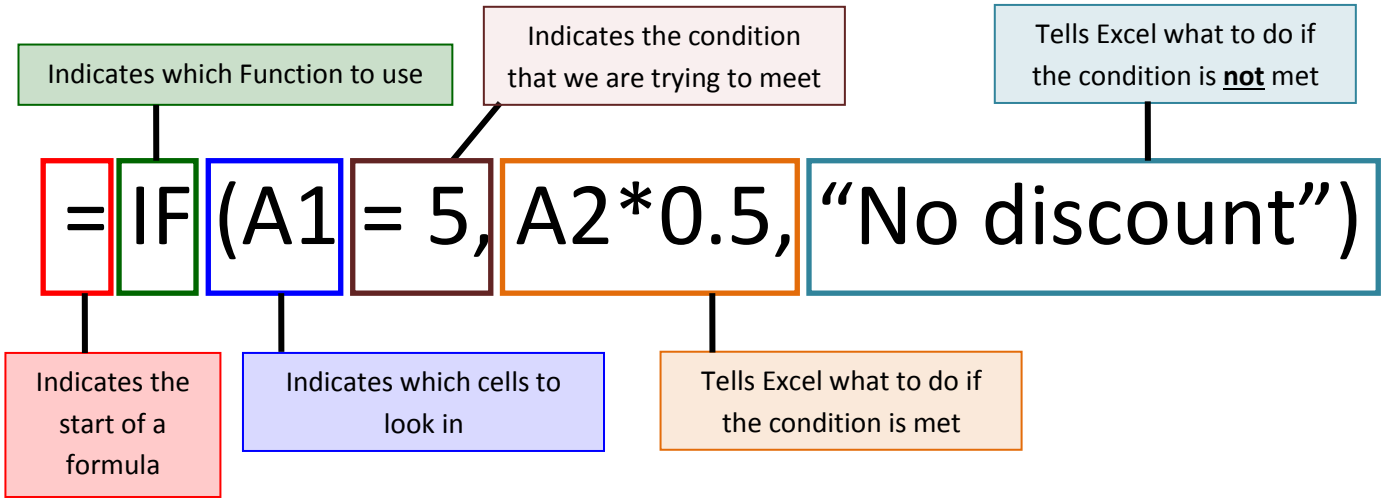
An **IF function** is made up of 3 parts:

- A condition
- What to do if the condition is met
- What to do if the condition is not met.

An example of an **IF function** is shown below:

=IF(A1=5, A2*0.5, "No discount")

Breakdown of the formula



The first part of the formula (the condition)

The first part of the formula **(A1=5)** is a **condition**. In this example it is **testing** to see if cell **A1 contains the number 5**.

The second part of the formula

The second part of the formula **(A2*0.5)** is telling Excel what to do if the above **condition is met**. In this case, the contents of cell **A2** will be **multiplied by 0.5**.

The third part of the formula

The third part of the formula **("No discount")** is telling Excel what to do if the above **condition is not met**. In this case, display the text **"No discount"**.

Task L

Open the file that you saved in **Data Analysis – Task I**.

Add a new label **'Category'** into cell **D2**.

Place formulae in cells **D3 to D21** to display **'Very experienced'** for employees with **ten years or more experience.**

For **every other employee** (those with less than ten years experience) the formula should display **'Not experienced'**.

Save the spreadsheet as **Data Analysis – Task L**

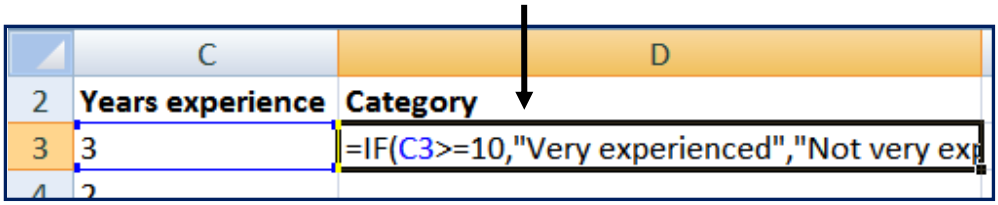
How to do it:

Open the file **Data Analysis – Task I**.

Place the cursor into cell **D2** and enter the label **'Category'**.

Place the cursor into cell **D3** and then enter the following formula:

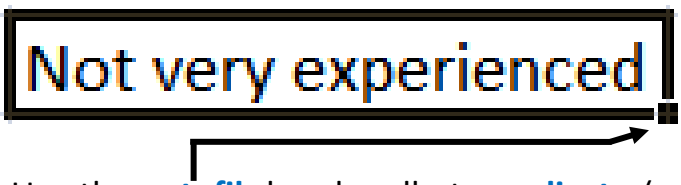
=IF(C3>=10,"Very experienced", "Not very experienced")



NOTE: The operator >= means more than or equal to….. in this case more than or equal to 10 (10 or more).

Do not use absolute cell referencing in this formula as the reference needs to change when you replicate the formula (Excel will reuse the same formula for the employee below and so on).

Press **<Enter>** to set the formula.



Use the **autofil** drag handle to **replicate** (copy) the **formula** all the way down to cell **D21**.

Your completed spreadsheet should now look like this:

Formula View

	C	D
2	**Years experience**	**Category**
3	3	=IF(C3>=10,"Very experienced","Not very experienced")
4	2	=IF(C4>=10,"Very experienced","Not very experienced")
5	12	=IF(C5>=10,"Very experienced","Not very experienced")
6	5	=IF(C6>=10,"Very experienced","Not very experienced")
7	7	=IF(C7>=10,"Very experienced","Not very experienced")
8	5	=IF(C8>=10,"Very experienced","Not very experienced")
9	6	=IF(C9>=10,"Very experienced","Not very experienced")
10	10	=IF(C10>=10,"Very experienced","Not very experienced")
11	14	=IF(C11>=10,"Very experienced","Not very experienced")
12	7	=IF(C12>=10,"Very experienced","Not very experienced")
13	3	=IF(C13>=10,"Very experienced","Not very experienced")
14	6	=IF(C14>=10,"Very experienced","Not very experienced")
15	9	=IF(C15>=10,"Very experienced","Not very experienced")
16	11	=IF(C16>=10,"Very experienced","Not very experienced")
17	10	=IF(C17>=10,"Very experienced","Not very experienced")
18	2	=IF(C18>=10,"Very experienced","Not very experienced")
19	1	=IF(C19>=10,"Very experienced","Not very experienced")
20	0.2	=IF(C20>=10,"Very experienced","Not very experienced")
21	2	=IF(C21>=10,"Very experienced","Not very experienced")

Value View

	A	B	C	D
1	Project 153			
2	**Name**	**Job**	**Years exp**	**Category**
3	Laila Aboli	Programmer	3	Not very experienced
4	Greg Mina	Programmer	2	Not very experienced
5	Sri Paryanti	Analyst	12	Very experienced
6	Bishen Patel	Sales	5	Not very experienced
7	Rupinder Singh	Engineer	7	Not very experienced
8	Sergio Gonzalez	Programmer	5	Not very experienced
9	Rupinder Vas	Sales	6	Not very experienced
10	Henri Ramos	Sales	10	Very experienced
11	John Mortlock	Programmer	14	Very experienced
12	Cameron Garnham	Analyst	7	Not very experienced
13	Brian Guthrie	Director	3	Not very experienced
14	Julia Frobisher	Engineer	6	Not very experienced
15	Dan McNevin	Programmer	9	Not very experienced
16	Patrick O'Malley	Engineer	11	Very experienced
17	Thirumalar Asokmani	Sales	10	Very experienced
18	Sean O'Byrne	Programmer	2	Not very experienced
19	Lea Cabusbusan	Programmer	1	Not very experienced
20	Brian O'Driscoll	Programmer	0.2	Not very experienced
21	Wim Van Hoffmann	Engineer	2	Not very experienced

Save the spreadsheet as **Data Analysis – Task L**
(Make sure it is saved as an Excel Workbook)

Activity 7

Open the file that you saved in **Activity 4**.

Add a new label **'New students'** into cell **F1**.

Place a formula in cells **F2 to F6** to display **'Add to this house'** if the number of students in the house is **fewer than six**.

If the number of students is **six or more**, the cell should display **'Full'**.

Save the document as **Data Analysis – Activity 7**.

NOTE:

NOW YOU SHOULD COMPLETE THE EXTRA IF FUNCTION TASKS AS PRACTICE

Nested IF Function

Nested functions contain a **function within another function**. They can handle **MORE than two outcomes** (A limitation of the normal IF Function)

Sometimes nested functions can contain several functions nested within each other.

Task M

Open the file that you saved in **Data Analysis – Task L**.

Change the formulae in cells **D3 to D21** to display **'Not experienced'** if they have **less than five years experience**.

Place formulae in cells **D3 to D21** to display **'Experienced'** for employees with **five years or more experience.**

For **every other employee** (those with ten or more year's experience) the formula should display **'Very experienced'**.

Save the spreadsheet as **Data Analysis – Task M**

How to do it:

For this task, **three conditions** exist:

- >=10 should display 'Very experienced' (more than or equal to 10)
- >=5 should display 'Experienced' (more than or equal to 5)
- <5 should display 'Not experienced' (Less than 5)

Place the cursor into cell **D3** and change the formula so that it becomes:

=IF(C3>=10,"Very experienced", IF(C3>=5,"Experienced", "Not experienced"))

NOTE: The second formula (highlighted in yellow), takes the place of the 'If the condition is not met' part of the first formula.

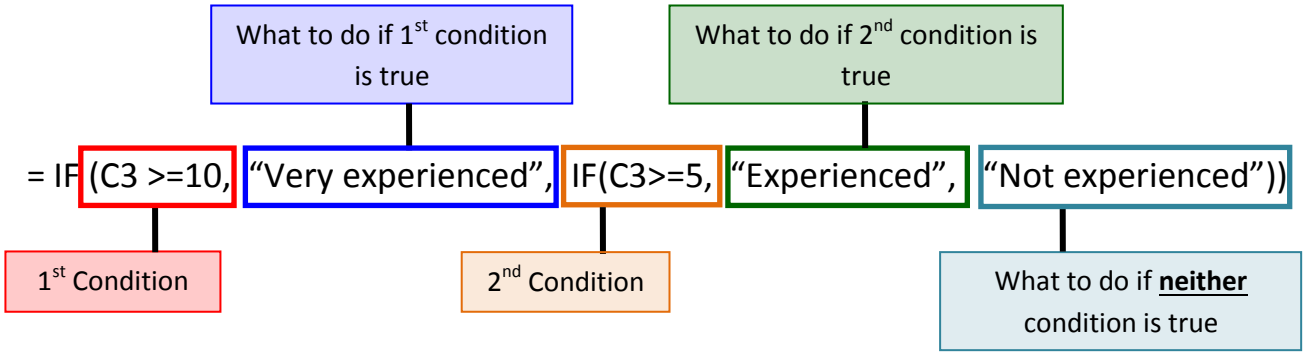
Notice how each condition has **one open** and **one close bracket**. The **open bracket** is

located **after each IF function** and the **close brackets** are located **at the end of the entire formula**.

You **MUST** work through the **conditions in order**. For example:

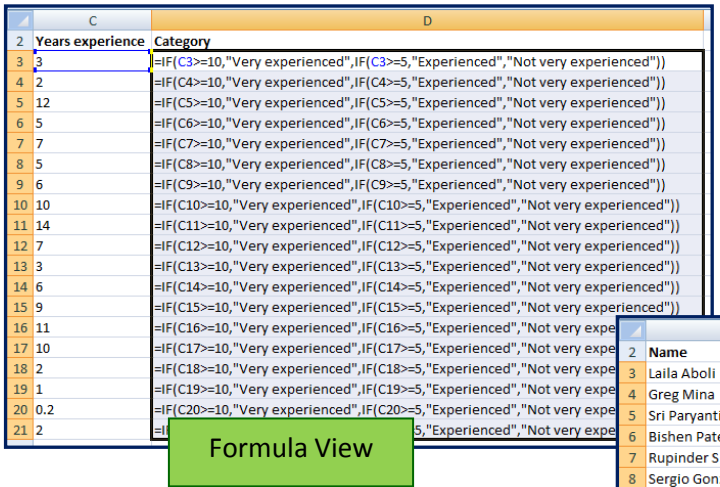
- Tell Excel what to do if C3 is more than or equal to 10……
- Tell Excel what to do if C3 is more than or equal to 5……..
- Tell Excel what to if none of the above are true (less than 5)

Breakdown of the formula



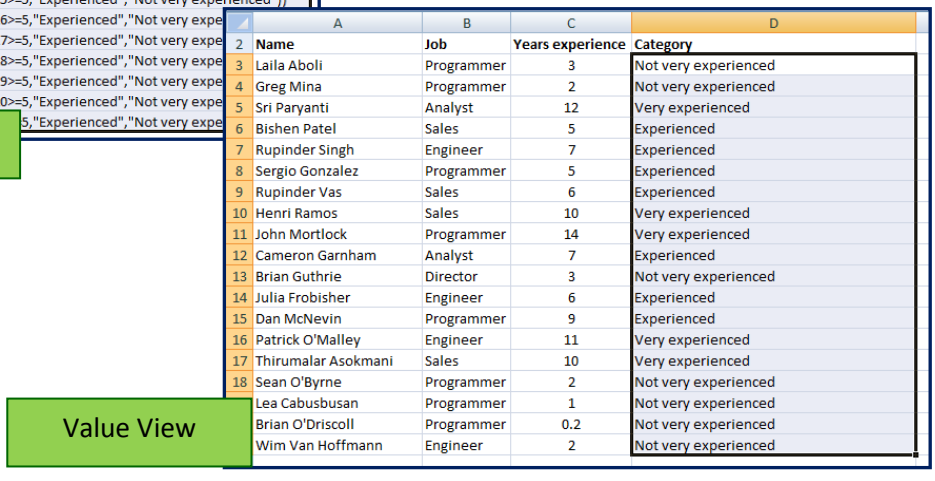
Use **Autofil** to replicate the formula down to **D21**.

Your completed spreadsheet should look like the images below:



	C	D
2	Years experience	Category
3	3	=IF(C3>=10,"Very experienced",IF(C3>=5,"Experienced","Not very experienced"))
4	2	=IF(C4>=10,"Very experienced",IF(C4>=5,"Experienced","Not very experienced"))
5	12	=IF(C5>=10,"Very experienced",IF(C5>=5,"Experienced","Not very experienced"))
6	5	=IF(C6>=10,"Very experienced",IF(C6>=5,"Experienced","Not very experienced"))
7	7	=IF(C7>=10,"Very experienced",IF(C7>=5,"Experienced","Not very experienced"))
8	5	=IF(C8>=10,"Very experienced",IF(C8>=5,"Experienced","Not very experienced"))
9	6	=IF(C9>=10,"Very experienced",IF(C9>=5,"Experienced","Not very experienced"))
10	10	=IF(C10>=10,"Very experienced",IF(C10>=5,"Experienced","Not very experienced"))
11	14	=IF(C11>=10,"Very experienced",IF(C11>=5,"Experienced","Not very experienced"))
12	7	=IF(C12>=10,"Very experienced",IF(C12>=5,"Experienced","Not very experienced"))
13	3	=IF(C13>=10,"Very experienced",IF(C13>=5,"Experienced","Not very experienced"))
14	6	=IF(C14>=10,"Very experienced",IF(C14>=5,"Experienced","Not very experienced"))
15	9	=IF(C15>=10,"Very experienced",IF(C15>=5,"Experienced","Not very experienced"))
16	11	=IF(C16>=10,"Very experienced",IF(C16>=5,"Experienced","Not very expe
17	10	=IF(C17>=10,"Very experienced",IF(C17>=5,"Experienced","Not very expe
18	2	=IF(C18>=10,"Very experienced",IF(C18>=5,"Experienced","Not very expe
19	1	=IF(C19>=10,"Very experienced",IF(C19>=5,"Experienced","Not very expe
20	0.2	=IF(C20>=10,"Very experienced",IF(C20>=5,"Experienced","Not very expe
21	2	=I 5,"Experienced","Not very expe

Save the spreadsheet as **Data Analysis – Task M**
(Make sure it is saved as an Excel Workbook)



	A	B	C	D
2	Name	Job	Years experience	Category
3	Laila Aboli	Programmer	3	Not very experienced
4	Greg Mina	Programmer	2	Not very experienced
5	Sri Paryanti	Analyst	12	Very experienced
6	Bishen Patel	Sales	5	Experienced
7	Rupinder Singh	Engineer	7	Experienced
8	Sergio Gonzalez	Programmer	5	Experienced
9	Rupinder Vas	Sales	6	Experienced
10	Henri Ramos	Sales	10	Very experienced
11	John Mortlock	Programmer	14	Very experienced
12	Cameron Garnham	Analyst	7	Experienced
13	Brian Guthrie	Director	3	Not very experienced
14	Julia Frobisher	Engineer	6	Experienced
15	Dan McNevin	Programmer	9	Experienced
16	Patrick O'Malley	Engineer	11	Very experienced
17	Thirumalar Asokmani	Sales	10	Very experienced
18	Sean O'Byrne	Programmer	2	Not very experienced
	Lea Cabusbusan	Programmer	1	Not very experienced
	Brian O'Driscoll	Programmer	0.2	Not very experienced
	Wim Van Hoffmann	Engineer	2	Not very experienced

Activity 8

Open the file that you saved in **Activity 7**.

Change the formulae in cells **F2** to **F6** to display:

- **'Add to this house'** if the **number of students** in this house is **fewer than six**.
- **'Ideal number'** if there are **between six and ten students**.
- **'Full'** if the number is **more than ten**.

Save the document as **Data Analysis – Activity 8**.

NOTE:

NOW YOU SHOULD COMPLETE THE EXTRA NESTED IF FUNCTION TASKS AS PRACTICE

IF AND Function

An **IF AND** function is similar to regular **IF** functions in that it allows a spreadsheet to **produce outcomes**.

The difference between **IF AND** and **regular IF Functions** is as follows:

- **Regular IF Functions** can perform just **one test** to determine the outcome
- **IF AND** can perform **two or more tests** that all determine the outcome.

Task N

Open the file called **Record.csv**.

Click in cell **E7** and enter a formula that will display the following:

- The word **"Yes"** if the athlete is **female** and their sprint time is **lower than the current world record for females**
- The word **"Yes"** if the athlete is **male** and their sprint time is **lower than the current world record for males**
- The word **"No"** for all other outcomes.

Replicate the formulae down to cell **E12**.

Save the spreadsheet as **Data Analysis – Task N**

For this task we have **two separate tests** that must **match to the criteria** set in the above question.

For females to break the record:

- Cell **C7** must contain **"Female"** (test 1)
- Cell **D7** must contain a **sprint time of less** than the **time held in cell C3** (test 2)

For males to break the record:

- Cell **C7** must contain **"Male"** (test 1)
- Cell **D7** must contain a **sprint time of less** than the **time held in cell C4** (test 2)

NOTE: Both cells C7 and D7 must match the criteria that we set or we must output "No" (as the correct records have not been broken).

How to do it:

Place the cursor into cell **E7** and enter the following formulae:

=IF(AND(C7="Female",D7<$C$3),"Yes",IF(AND(C7="Male",D7<$C$4),"Yes","No"))

> **NOTE:** The **IF(AND (test 1, test2)** part of the formulae allows us to perform two different tests

Breakdown of the formula

Two tests – one to see if athlete is **female** and one to see if the female **record has been broken**

Two tests – one to see if athlete is **male** and one to see if the male **record has been broken**

What to output if **no exact matches** to the tests are found

= IF(AND(C7="Female",D7<$C$3) ,"Yes" , IF(AND(C7="Male",D7<$C$4) ,"Yes" ,"No"))

Allows us to test 2 or more cells

What to output if cell C7 is female **AND** cell D7 is less than the contents of cell C3

What to output if cell C7 is male **AND** cell D7 is less than the contents of cell C4

Use **Autofil** to replicate the formula down to **E12**.

Your completed spreadsheet should look like the images below:

	A	B	C	D	E
6		**Name**	**Gender**	**Sprint Time**	**Record Broken?**
7		Mike	Male	11.03	=IF(AND(C7="Female",D7<$C$3),"Yes",IF(AND(C7="Male",D7<$C$4),"Yes","No"))
8		Salma	Female	10.49	=IF(AND(C8="Female",D8<$C$3),"Yes",IF(AND(C8="Male",D8<$C$4),"Yes","No"))
9		Rachel	Female	10.23	=IF(AND(C9="Female",D9<$C$3),"Yes",IF(AND(C9="Male",D9<$C$4),"Yes","No"))
10		Jim	Male	9.6	=IF(AND(C10="Female",D10<$C$3),"Yes",IF(AND(C10="Male",D10<$C$4),"Yes","No"))
11		Peter	Male	9.56	=IF(AND(C11="Female",D11<$C$3),"Yes",IF(AND(C11="Male",D11<$C$4),"Yes","No"))
12		Florence	Female	10.78	=IF(AND(C12="Female",D12<$C$3),"Yes",IF(AND(C12="Male",D12<$C$4),"Yes","No"))

Formula View

	A	B	C	D	E
6		**Name**	**Gender**	**Sprint Time**	**Record Broken?**
7		Mike	Male	11.03	No
8		Salma	Female	10.49	No
9		Rachel	Female	10.23	Yes
10		Jim	Male	9.6	No
11		Peter	Male	9.56	Yes
12		Florence	Female	10.78	No

Value View

Using Lookups

The term 'Look up' as used in the examinations, means to look up (search) from a list.

There are two main LOOKUP functions that can be used within Excel. These are:

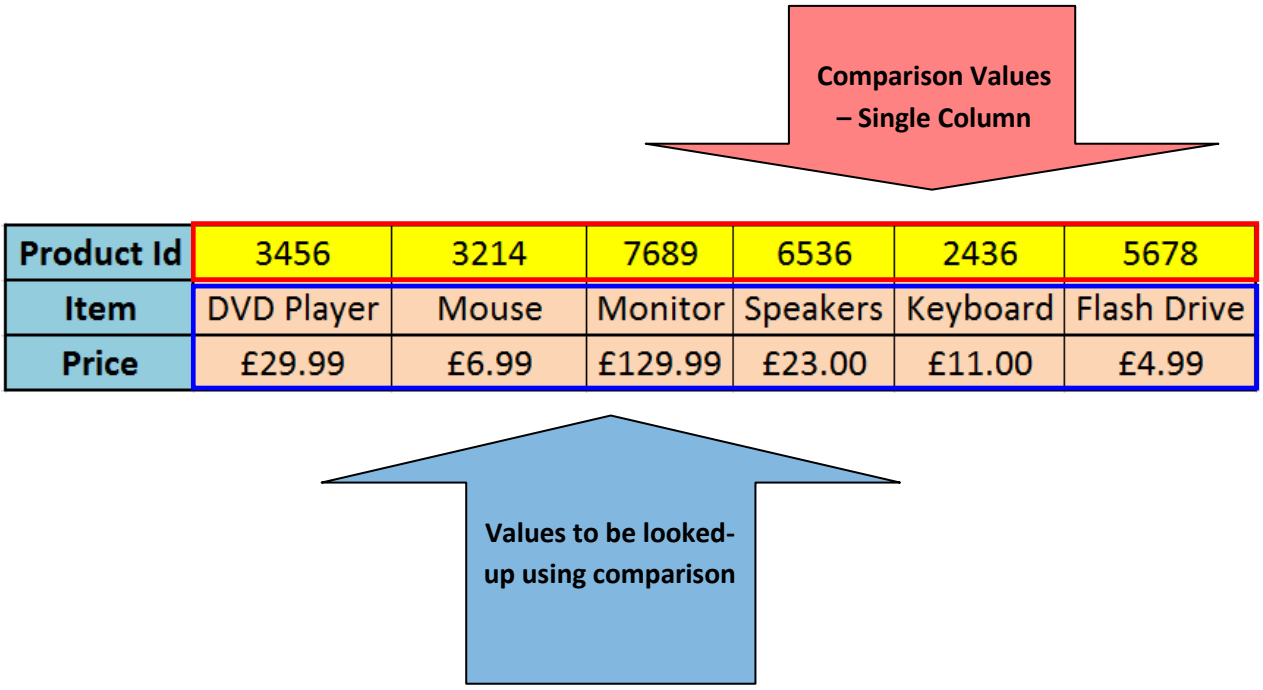
- HLOOKUP
- VLOOKUP

HLOOKUP (Horizontal Lookup)

HLOOKUP is a function that performs a **horizontal look-up of data**.

HLOOKUP should be used when the **values** you wish to **compare your data with** are **stored in a single row**.

The **values to be looked up** are stored in the **rows below the comparison values**. This is shown in the diagram below:



Product Id	3456	3214	7689	6536	2436	5678
Item	DVD Player	Mouse	Monitor	Speakers	Keyboard	Flash Drive
Price	£29.99	£6.99	£129.99	£23.00	£11.00	£4.99

Task O

Open the file **JOBS.CSV**.

Insert formulae in the Description column to look up and display the JobTitle using the JobCode as the look-up value:

Save the spreadsheet as **Data Analysis – Task O**

How to do it:

Open the **JOBS.CSV** file and place the cursor in cell **C6**. Enter the formula:

=HLOOKUP(B6, $B$2:$H$3, 2, FALSE)

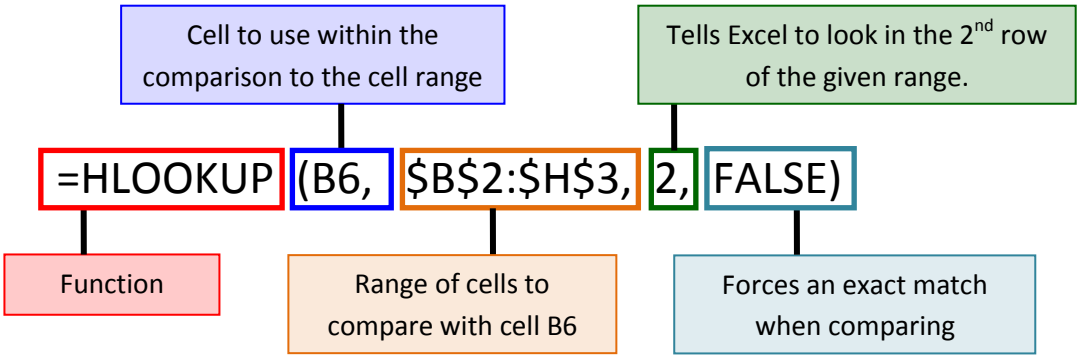
This formula will look up and **compare the contents of cell B6** with the **contents of each cell in the top row of the range B2 to H3**.

When it finds a match, it will take the value or label stored in the second row which is directly under the matched cell.

NOTE: The 2 at the end of the formula tells Excel to look in the second row of the given range.

The FALSE part forces Excel to match the values exactly.

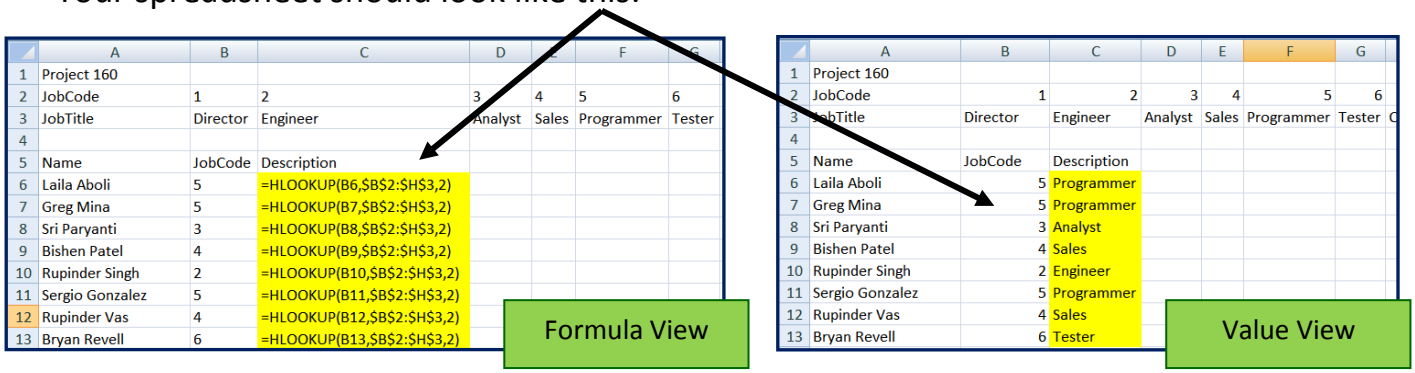
Breakdown of the formula



NOTE: The number 2 near the end of the formula could point to any row within the range, if it exists. (3 for example)

Replicate the formula (using autofil) down to cell **C27**.

Your spreadsheet should look like this:



	A	B	C	D	E	F	G
1	Project 160						
2	JobCode	1	2	3	4	5	6
3	JobTitle	Director	Engineer	Analyst	Sales	Programmer	Tester
4							
5	Name	JobCode	Description				
6	Laila Aboli	5	=HLOOKUP(B6,$B$2:$H$3,2)				
7	Greg Mina	5	=HLOOKUP(B7,$B$2:$H$3,2)				
8	Sri Paryanti	3	=HLOOKUP(B8,$B$2:$H$3,2)				
9	Bishen Patel	4	=HLOOKUP(B9,$B$2:$H$3,2)				
10	Rupinder Singh	2	=HLOOKUP(B10,$B$2:$H$3,2)				
11	Sergio Gonzalez	5	=HLOOKUP(B11,$B$2:$H$3,2)				
12	Rupinder Vas	4	=HLOOKUP(B12,$B$2:$H$3,2)				
13	Bryan Revell	6	=HLOOKUP(B13,$B$2:$H$3,2)				

Formula View

	A	B	C	D	E	F	G
1	Project 160						
2	JobCode	1	2	3	4	5	6
3	JobTitle	Director	Engineer	Analyst	Sales	Programmer	Tester
4							
5	Name	JobCode	Description				
6	Laila Aboli	5	Programmer				
7	Greg Mina	5	Programmer				
8	Sri Paryanti	3	Analyst				
9	Bishen Patel	4	Sales				
10	Rupinder Singh	2	Engineer				
11	Sergio Gonzalez	5	Programmer				
12	Rupinder Vas	4	Sales				
13	Bryan Revell	6	Tester				

Value View

Save the spreadsheet as **Data Analysis – Task O**
(Make sure it is saved as an Excel Workbook)

NOTE:

NOW YOU SHOULD COMPLETE THE EXTRA HLOOKUP FUNCTION TASKS AS PRACTICE

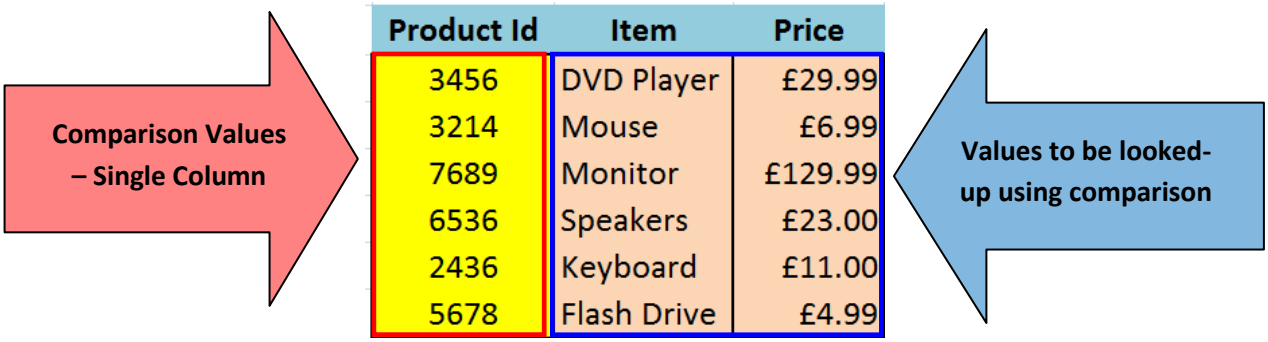
VLOOKUP (Vertical Lookup)

VLOOKUP is a function that performs a **vertical look-up of data**.

VLOOKUP should be used when the **values** you wish to **compare your data with** are **stored in a single column**.

The **values to be looked up** are stored in the **columns to the right of the comparison** values. This is shown in the diagram below:

NOTE: The look-up values can be stored in the same spreadsheet file or in a different spreadsheet file.



Product Id	Item	Price
3456	DVD Player	£29.99
3214	Mouse	£6.99
7689	Monitor	£129.99
6536	Speakers	£23.00
2436	Keyboard	£11.00
5678	Flash Drive	£4.99

Using VLOOKUP within the same spreadsheet file

Task P

Open the file **TASKS.CSV**.

Insert formulae into the **Current Task column** to look up and display the **current tasks** for each of the ICT experts in our company using the **TaskCode** for the **look-up value** and the **VLOOKUP table**.

Use both **absolute** and **relative cell referencing** within the formulae.

Save the spreadsheet as **Data Analysis – Task P**

How to do it:

Open the file **TASKS.CSV** and place the cursor into cell **C3**. Enter the formula:

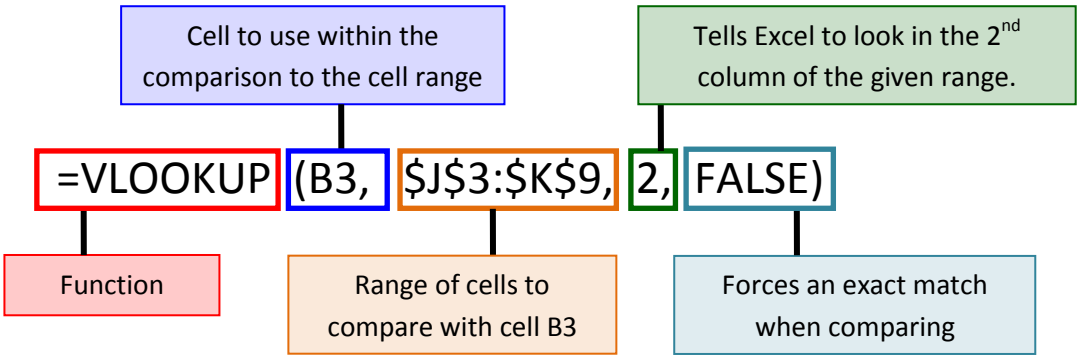
=VLOOKUP(B3, $J$3:$K$9, 2, FALSE)

This formula will look up and **compare the contents of cell B3** with the **contents of each cell in the 1st column of the range J3 to K9**.

When it finds a match, it will take the value or label stored in the second column which is to the right of the matched data.

> **NOTE: The 2 near the end of the formula tells Excel to look in the second column of the given range.**
>
> **The FALSE part forces Excel to match the values exactly.**

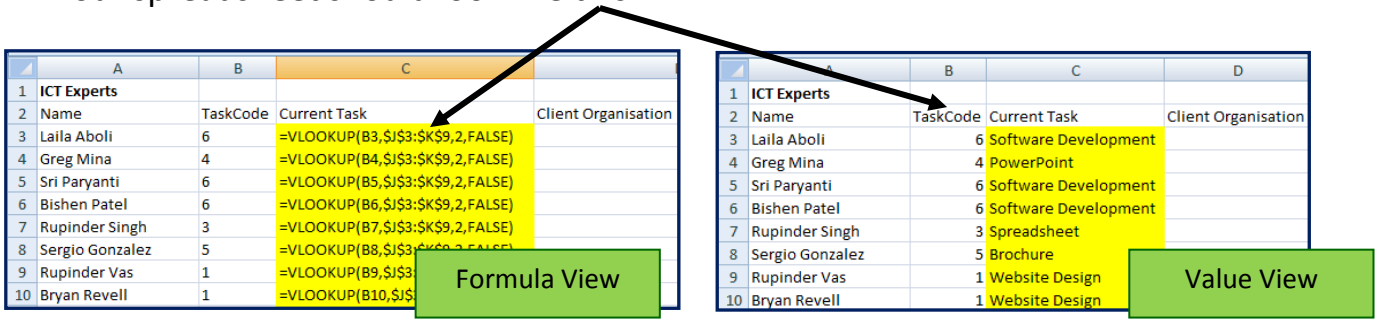
Breakdown of the formula



> **NOTE: The number 2 near the end of the formula could point to any column within the range, if it exists. (3 for example)**

Replicate the formula (using autofil) down to cell **C24**.

Your spreadsheet should look like this:



	A	B	C	
1	ICT Experts			
2	Name	TaskCode	Current Task	Client Organisation
3	Laila Aboli	6	=VLOOKUP(B3,$J$3:$K$9,2,FALSE)	
4	Greg Mina	4	=VLOOKUP(B4,$J$3:$K$9,2,FALSE)	
5	Sri Paryanti	6	=VLOOKUP(B5,$J$3:$K$9,2,FALSE)	
6	Bishen Patel	6	=VLOOKUP(B6,$J$3:$K$9,2,FALSE)	
7	Rupinder Singh	3	=VLOOKUP(B7,$J$3:$K$9,2,FALSE)	
8	Sergio Gonzalez	5	=VLOOKUP(B8,$J$3:$K$9,2,FALSE)	
9	Rupinder Vas	1	=VLOOKUP(B9,$J$3:$K$9,2,FALSE)	
10	Bryan Revell	1	=VLOOKUP(B10,$J$3:$K$9,2,FALSE)	

	A	B	C	D
1	ICT Experts			
2	Name	TaskCode	Current Task	Client Organisation
3	Laila Aboli	6	Software Development	
4	Greg Mina	4	PowerPoint	
5	Sri Paryanti	6	Software Development	
6	Bishen Patel	6	Software Development	
7	Rupinder Singh	3	Spreadsheet	
8	Sergio Gonzalez	5	Brochure	
9	Rupinder Vas	1	Website Design	
10	Bryan Revell	1	Website Design	

Save the spreadsheet as **Data Analysis – Task P**

Using VLOOKUP between different spreadsheet files (Also applies to HLOOKUP)

Task Q

Open the file **TASK M**.

Insert formulae in the **Client Organisation column** to look up and display the **client**, using the **TaskCode** for the **look-up value** and the file **CLIENT.CSV**.

Use both **absolute** and **relative cell referencing** within the formulae.

Save the spreadsheet as **Data Analysis – Task Q**

How to do it:

Open up the file **Task M** **AND** the file **CLIENT.CSV**.

> **NOTE: It is essential that both spreadsheets to be used in look-up are open**

Making sure that you are within the file **Task M**, place the cursor into cell **D3**.

Enter the following formula:

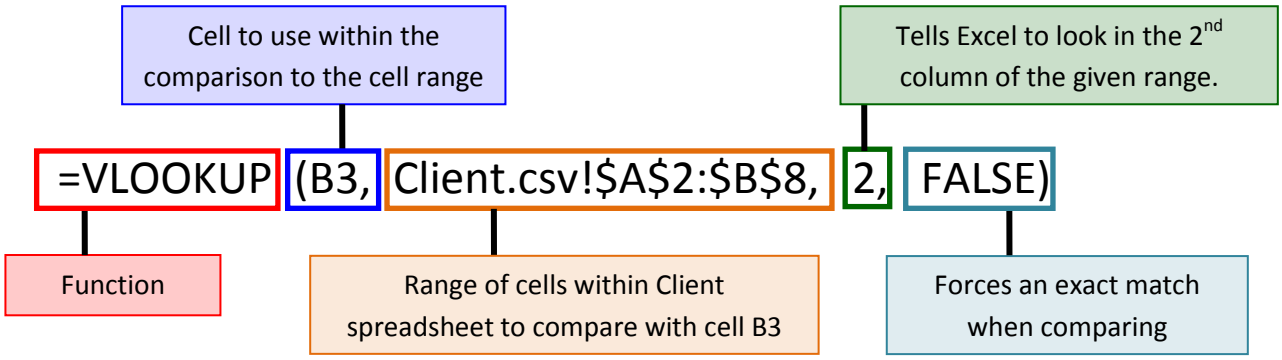
=VLOOKUP(B3, Client.csv!$A$2:$B$8, 2, FALSE)

This formula will look up and **compare the contents of cell B3** with the **contents of each cell in the 1st column of the range A2 to B8 within the file CLIENT.CSV**.

NOTE: The yellow highlighted section of this formula can be added by either:

- **Typing it in manually (hard to remember)**
- **Clicking your mouse cursor into the file and highlighting all cells in both columns of the lookup table.**

Breakdown of the formula



REMEMBER: **The 2 near the end of the formula tells Excel to look in the second column of the given range.**

The FALSE part forces Excel to only display the values if they match exactly.

Replicate the formula (using autofil) down to cell **D24**.

Your spreadsheet should look like this:

	A	B	C	D
1	**ICT Experts**			
2	Name	TaskCode	Current Task	Client Organisation
3	Laila Aboli	6	=VLOOKUP(B3,$J$3:$K$9,2,FALSE)	=VLOOKUP(B3,Client.CSV!$A$2:$B$8,2,FALSE)
4	Greg Mina	4	=VLOOKUP(B4,$J$3:$K$9,2,FALSE)	=VLOOKUP(B4,Client.CSV!$A$2:$B$8,2,FALSE)
5	Sri Paryanti	6	=VLOOKUP(B5,$J$3:$K$9,2,FALSE)	=VLOOKUP(B5,Client.CSV!$A$2:$B$8,2,FALSE)
6	Bishen Patel	6	=VLOOKUP(B6,$J$3:$K$9,2,FALSE)	=VLOOKUP(B6,Client.CSV!$A$2:$B$8,2,FALSE)
7	Rupinder Singh	3	=VLOOKUP(B7,$J$3:$K$9,2,FALSE)	=VLOOKUP(B7,Client.CSV!$A$2:$B$8,2,FALSE)
8	Sergio Gonzalez	5	=VLOOKUP(B8,$J$3:$K$9,2,FALSE)	=VLOOKUP(B8,Client.CSV!$A$2:$B$8,2,FALSE)
9	Rupinder Vas	1	=VLOOKUP(B9,$J$3:$K$9,2,FALSE)	=VLOOKUP(B9,Client.CSV!$A$2:$B$8,2,
10	Bryan Revell	1	=VLOOKUP(B10,$J$3:$K$9,2,FALSE)	=VLOOKUP(B10,Client.CSV!$A$2:$B$8,2

Formula View

Save the spreadsheet as **Data Analysis – Task Q**

	A	B	C	D
1	**ICT Experts**			
2	Name	TaskCode	Current Task	Client Organisation
3	Laila Aboli	6	Software Development	LGY
4	Greg Mina	4	PowerPoint	Avricom
5	Sri Paryanti	6	Software Development	LGY
6	Bishen Patel	6	Software Development	LGY
7	Rupinder Singh	3	Spreadsheet	Hothouse Design
8	Sergio Gonzalez	5	Brochure	Binnaccount
9	Rupinder Vas			Rootrainer
10	Bryan Revell			Rootrainer

Value View

Activity 9

Open the file **TUTORS.CSV**.

Insert formulae in the **Tutor Name column** to look up the **tutor's name** by **matching the tutor's initials** to the file **TEACHERS.CSV**.

Insert formulae in the **Room Number column** to look up the **room number** by **matching the tutor's initials** to the file **ROOMS.CSV**. **(This formula is HLookup not VLookup)**

Make sure that you use appropriate **absolute** and **relative cell referencing**.

Save the document as **Data Analysis – Activity 9**.

NOTE:

NOW YOU SHOULD COMPLETE THE EXTRA VLOOKUP FUNCTION TASKS AS PRACTICE

14.7 – Interrogating Data (Using Filters)

Interrogating data in Excel refers to the task of **extracting data** that meets **certain criteria.**

We can use the **Filter tool** to accomplish this.

Task R

Open the file **TASK Q**.

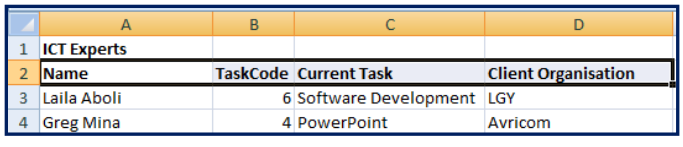
Interrogate the data to search for the **employees** who are currently working on jobs for **Binnaccount**.

Save the spreadsheet as **Data Analysis – Task R**

How to do it:

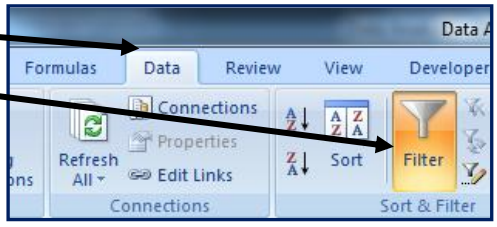
Open up the file **Task Q**.

Highlight the data headings **(cells A2 to D2)**



	A	B	C	D
1	ICT Experts			
2	Name	TaskCode	Current Task	Client Organisation
3	Laila Aboli	6	Software Development	LGY
4	Greg Mina	4	PowerPoint	Avricom

Select the **Data tab** and click on the **Filter** icon.

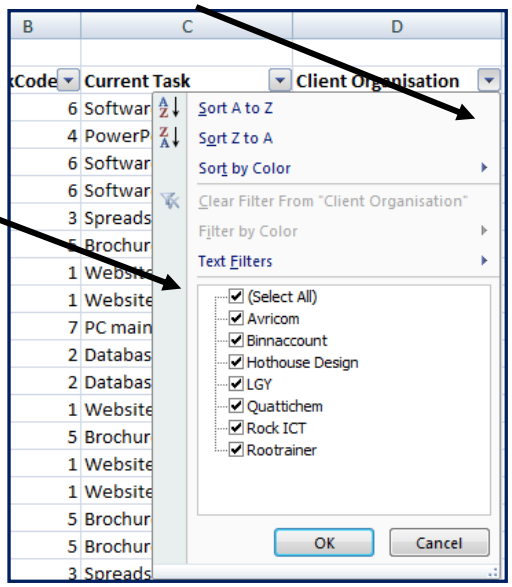


This will display an **arrow** in the **top corner** of each **column**.

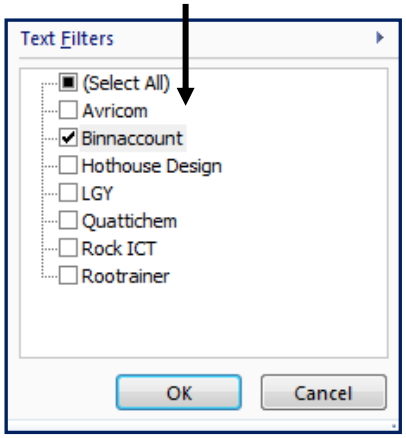
	A	B	C	D
1	ICT Experts			
2	Name	TaskCode	Current Task	Client Organisation
3	Laila Aboli	6	Software Development	LGY
4	Greg Mina	4	PowerPoint	Avricom
5	Sri Paryanti	6	Software Development	LGY
6	Bishen Patel	6	Software Development	LGY

For this task we need to extract the data of **employees who are currently working for the Binnaccount organisation**. To do this we need to **click the arrow** in the **Client Organisation cell**.

When you do this a **drop-down menu** should appear.



In the **Text Filters** section of the menu, remove the ticks from every box except the Binnaccount box.



Click **OK**.

This will display on the data of the employees who are currently working for the Binnaccount organisation. All other employees are hidden.

	A	B	C	D
1	ICT Experts			
2	Name	TaskCode	Current Task	Client Organisation
8	Sergio Gonzalez	5	Brochure	Binnaccount
15	Julia Frobisher	5	Brochure	Binnaccount
18	Patrick O'Malley	5	Brochure	Binnaccount
19	Thirumalar Asokmani	5	Brochure	Binnaccount
24	Wim Van Hoffmann	5	Brochure	Binnaccount

NOTE: The same method can be used to select more than one company from the list.

14.7b – Interrogating Data using Number Filters

You can also interrogate data using numbers as the criteria.

Task S

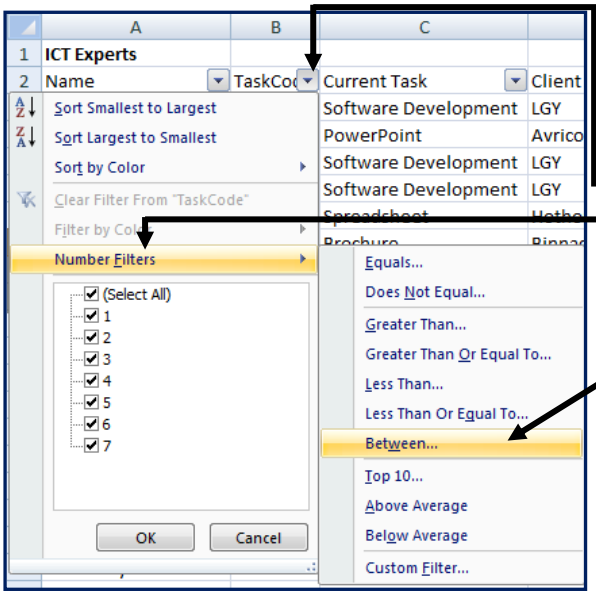
Open the file **TASK Q**.

Interrogate the data to search for the **employees** where the **task code in between three and six**.

Save the spreadsheet as **Data Analysis – Task S**

How to do it:

Open up the file **Task Q**.

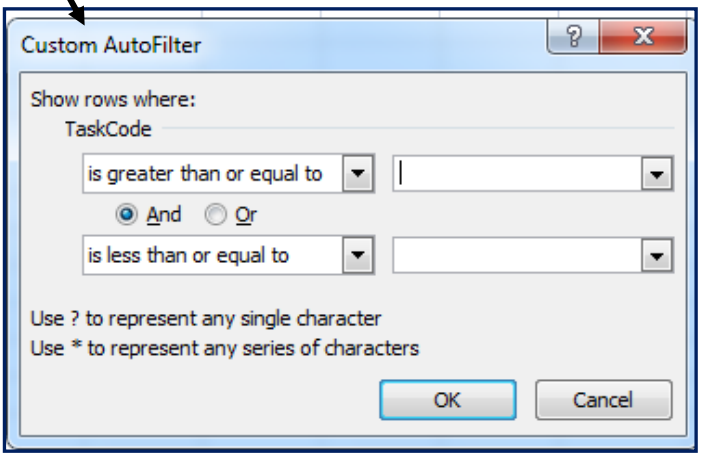


Highlight the data headings **(cells A2 to D2)**.

Click on the **Filter icon** and click the arrow in the **Task Code** column.

Select the **Number Filters** option and then **Between…**

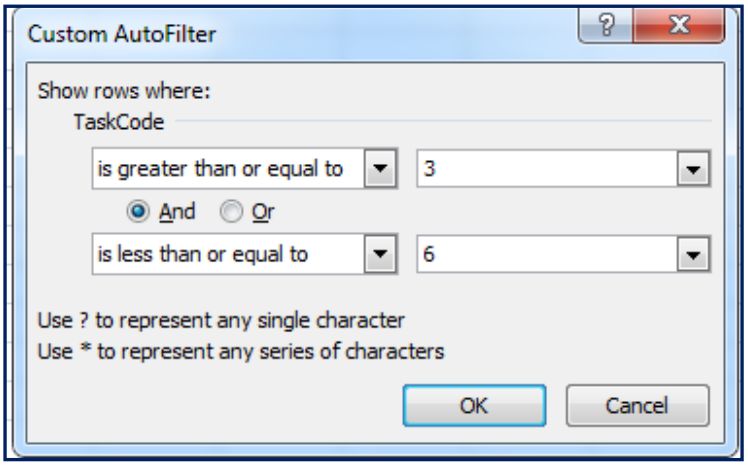
This gives you access to the **Custom AutoFilter** window.



NOTE: In a 'Between Filter', Excel automatically adds the criteria for you.

(is greater than or equal to) AND (is less than or equal to)

This allows you to set the search criteria to **greater than or equal to 3 AND less than or equal to 6**.



NOTE:

You can select and use the following number filters in the same way:

- **Equal to**
- **Not equal to**
- **Less than**
- **Greater than**
- **etc**

Your interrogated (filtered) spreadsheet should now display only the data with Task Codes from 3 to 6:

	A	B	C	D
1	ICT Experts			
2	Name	TaskCode	Current Task	Client Organisation
3	Laila Aboli	6	Software Development	LGY
4	Greg Mina	4	PowerPoint	Avricom
5	Sri Paryanti	6	Software Development	LGY
6	Bishen Patel	6	Software Development	LGY
7	Rupinder Singh	3	Spreadsheet	Hothouse Design
8	Sergio Gonzalez	5	Brochure	Binnaccount
15	Julia Frobisher	5	Brochure	Binnaccount
18	Patrick O'Malley	5	Brochure	Binnaccount
19	Thirumalar Asokmani	5	Brochure	Binnaccount
20	Sean O'Byrne	3	Spreadsheet	Hothouse Design
23	Brian O'Driscoll	3	Spreadsheet	Hothouse Design
24	Wim Van Hoffmann	5	Brochure	Binnaccount

Save the spreadsheet as **Data Analysis – Task S**

Activity 10

Open the file you saved in **Activity 9**.

Select from the **all of the data**:

- All the students with a **tutor** called **Chris Scott**
- All the students who will be using the **rooms numbered between 22 and 74**
- All the students, **except Kiah and Hartati**, with a **tutor called Kate Morrissey or Mike Arnott**.

NOTE: You MUST clear each filter before you start the next.

Save the document as **Data Analysis – Activity 10**.

NOTE:

The following skills are described on videos found on the ICT Lounge website in IGCSE Section 14: http://www.ictlounge.com/html/dataanalysis.htm

- **Creating Charts / Graphs**
- **Hiding and Showing Columns / Rows**
- **Printing spreadsheets**