IGCSE ICT – SECTION 14
DATA ANALYSIS

MICROSOFT EXCEL
STEP BY STEP GUIDE

Mark Nicholls
ICT Lounge
Data Analysis – Self Study Guide

Contents

Learning Outcomes ...................................................................................................................... Page 3
What is a Data Model? .................................................................................................................. Page 4
Spreadsheet Basics...................................................................................................................... Page 4 – 6
Resizing Column Widths ............................................................................................................. Page 6
Autofil ........................................................................................................................................ Page 7
Absolute Cell Reference ............................................................................................................. Page 8
Printing Spreadsheet Values ..................................................................................................... Page 9
Printing Spreadsheet Formulae .................................................................................................. Page 10
Basis Formulae (Operators) ....................................................................................................... Page 11 – 13
Naming Cells and Cell Ranges .................................................................................................. Page 14 – 15
Introduction to Functions ......................................................................................................... Page 16
   SUM Function .......................................................................................................................... Page 16 – 17
   AVERAGE Function ............................................................................................................... Page 18
   MAX Function ......................................................................................................................... Page 18
   MIN Function ......................................................................................................................... Page 19
Alternative to Typing in Functions ............................................................................................. Page 19 – 20
Activity 1 .................................................................................................................................... Page 20
   INT Function .......................................................................................................................... Page 21 – 22
   ROUND Function .................................................................................................................. Page 22 – 23
Activity 2 .................................................................................................................................... Page 23
   COUNT Function ................................................................................................................... Page 24
   COUNTA Function ................................................................................................................ Page 25
Activity 3 .................................................................................................................................... Page 26
   COUNTIF Function ................................................................................................................ Page 26 – 28
Activity 4 .................................................................................................................................... Page 29
   SUMIF Function ...................................................................................................................... Page 30 – 31
Activity 5 .................................................................................................................................... Page 32
   SUMIF Function with NOT criteria ...................................................................................... Page 33 – 34
   COUNTIF Function with NOT criteria ................................................................................. Page 35 – 36
Activity 6 .................................................................................................................................... Page 36
   IF Function ............................................................................................................................. Page 37 – 39
Activity 7 ........................................................................................................................................ Page 39
  NESTED IF Function .............................................................................................................. Page 40 – 41
Activity 8 ..................................................................................................................................... Page 42
  IF AND Function .................................................................................................................... Page 43 – 44
Using Lookups .......................................................................................................................... Page 45
  HLOOKUP Function ............................................................................................................... Page 45 – 47
  VLOOKUP Function ............................................................................................................... Page 48 – 49
  LOOKUP Function between 2 Spreadsheets ........................................................................... Page 50 – 51
Activity 9 ..................................................................................................................................... Page 52
  Interrogating Data (Using Filters) .......................................................................................... Page 52 - 53
    Numbered Filters .................................................................................................................. Page 54 - 55
Activity 10 .................................................................................................................................. Page 56
  Finding help for Graphs and Charts ...................................................................................... Page 56
  Finding help for Hiding/Showing Columns and Rows ............................................................... Page 56
  Finding help for Printing Gridlines and Areas of Spreadsheets ................................................. Page 56
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.

![Spreadsheet Layout](image)

**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**

- Indicates which cells to look in
- Indicates which calculation (operator) to use
- Indicates the start of a 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.

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.

**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:

\[ =A3 \times \$A\$1 \]

**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 Autofill 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.
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 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 B1 divided 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. **Addition Calculation:**
   
   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: \( \frac{B1}{B2} \)

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.

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:

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.

How to do it:

**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**

**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.

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.

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.

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

**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**.

**Breakdown of the formula**

$$=\text{SUM}(B4:B8)$$

- **Indicates the start of a formula**
- **Indicates which Function to use**
- **Indicates which cells to look in**

**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:

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

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

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.
Section 14: Data Analysis

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.

Your final spreadsheet should look like this:

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**.

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.
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.

**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**

- Indicates the start of a formula
- Indicates which Function to use
- Indicates which cells to look in
- Indicates how many decimal places to round to
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.

<table>
<thead>
<tr>
<th>Function</th>
<th>Result of Rounding</th>
<th>What it does</th>
</tr>
</thead>
<tbody>
<tr>
<td><code>=ROUND(A1,2)</code></td>
<td>62.55</td>
<td>Rounds the contents of A1 to two decimal places.</td>
</tr>
<tr>
<td><code>=ROUND(A1,1)</code></td>
<td>62.6</td>
<td>Rounds the contents of A1 to one decimal place. The digit ‘5’ has forced the previous figure to be rounded up.</td>
</tr>
<tr>
<td><code>=ROUND(A1,0)</code></td>
<td>63</td>
<td>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.</td>
</tr>
<tr>
<td><code>=ROUND(A1,-1)</code></td>
<td>60</td>
<td>Rounds the contents of A1 to the nearest 10.</td>
</tr>
<tr>
<td><code>=ROUND(A1,-2)</code></td>
<td>100</td>
<td>Rounds the contents of cell A1 to the nearest 100.</td>
</tr>
</tbody>
</table>

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

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 values 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:

\[ =\text{COUNTA}(A2:A19) - \text{COUNT}(A2:A19) \]

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.

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:

<table>
<thead>
<tr>
<th>Function</th>
<th>What it does</th>
</tr>
</thead>
<tbody>
<tr>
<td>=COUNTIF($B$3:$B$21, “Director”)</td>
<td>Counts the number of cells in the range B3 to B21 that contain the word ‘Director’.</td>
</tr>
<tr>
<td>=COUNTIF(Job, “Director”)</td>
<td>Counts the number of cells in the named range ‘Job’ (B3 to B21) that contain the word ‘Director’.</td>
</tr>
<tr>
<td>=COUNTIF($B$3:$B$21, A24)</td>
<td>Counts the number of cells in the range B3 to B21 that contain the same text as the contents of cell A24.</td>
</tr>
<tr>
<td>=COUNTIF(Job, A24)</td>
<td>Counts the number of cells in the named range ‘Job’ (B3 to B21) that contain the same text as the contents of cell A24.</td>
</tr>
</tbody>
</table>

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

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.

**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:

\[=\text{COUNTIF}(\$C\!3:\$C\!21, "<5")\]

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**.

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

Enter the following formula:

\[=\text{COUNTIF}(\$C\!3:\$C\!21, "\geq 10")\]

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**.

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

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.
Breakdown of the formula

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

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 autofill 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**

- Indicates which Function to use
- Indicates the start of a formula
- Tells Excel to exclude cells containing NT from the SUMIF calculation
- Indicates which cells to look in
=SUMIF(D8:I8, "<>NT")

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

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 autofill 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**

- Indicates which Function to use
- Tells Excel to exclude cells containing NT from the SUMIF calculation
- Indicates which cells to look in
- Indicates the start of a formula
Use the autofil drag handle to replicate (copy) the formula all the way down to cell K80.

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 autofill drag handle to replicate (copy) the formula all the way down to cell D21.
Your completed spreadsheet should now look like this:

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**

```
= IF (C3 >=10, “Very experienced”, IF(C3>=5, “Experienced”, “Not experienced”))
```

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

Your completed spreadsheet should look like the images below:

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

```excel
=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
  
  \[
  \text{IF(AND(C7="Female",D7<$C$3),"Yes",IF(AND(C7="Male",D7<$C$4),"Yes","No"))}
  \]

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

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

- **Use Autofill to replicate the formula down to E12.**

Your completed spreadsheet should look like the images below:
**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:

<table>
<thead>
<tr>
<th>Product Id</th>
<th>3456</th>
<th>3214</th>
<th>7689</th>
<th>6536</th>
<th>2436</th>
<th>5678</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
<td>DVD Player</td>
<td>Mouse</td>
<td>Monitor</td>
<td>Speakers</td>
<td>Keyboard</td>
<td>Flash Drive</td>
</tr>
<tr>
<td>Price</td>
<td>£29.99</td>
<td>£6.99</td>
<td>£129.99</td>
<td>£23.00</td>
<td>£11.00</td>
<td>£4.99</td>
</tr>
</tbody>
</table>
**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**

- **Cell to use within the comparison to the cell range**
- **Range of cells to compare with cell B6**
- **Tells Excel to look in the 2nd row of the given range.**
- **Forces an exact match when comparing**

**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:

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:

![Diagram showing VLOOKUP concept](image)

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

### 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

- **Function**: VLOOKUP
- **Cell to use within the comparison to the cell range**: B3
- **Range of cells to compare with cell B3**: $J$3:$K$9
- **Tells Excel to look in the 2\textsuperscript{nd} column of the given range**: 2
- **Forces an exact match when comparing**: FALSE

**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 autofill) down to cell C24.

Your spreadsheet should look like this:

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:

\[ =VLOOKUP(B3, \text{Client.csv!$A$2:$B$8}, 2, \text{FALSE}) \]

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

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

- **Function**: VLOOKUP
- **Cell to use within the comparison to the cell range**: B3
- **Range of cells within Client spreadsheet to compare with cell B3**: Client.csv!$A$2:$B$8
- **Tells Excel to look in the 2nd column of the given range**: 2
- **Forces an exact match when comparing**: FALSE

**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 autofill) down to cell D24.

Your spreadsheet should look like this:

Save the spreadsheet as **Data Analysis – Task Q**
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)
Section 14: Data Analysis

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

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

<table>
<thead>
<tr>
<th></th>
<th>ICT Experts</th>
<th>TaskCode</th>
<th>Current Task</th>
<th>Client Organisation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Name</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Laila Aboli</td>
<td></td>
<td>6 Software Development</td>
<td>LGY</td>
</tr>
<tr>
<td>4</td>
<td>Greg Mina</td>
<td>4</td>
<td>PowerPoint</td>
<td>Avricom</td>
</tr>
<tr>
<td>5</td>
<td>Sri Paryanti</td>
<td>6</td>
<td>Software Development</td>
<td>LGY</td>
</tr>
<tr>
<td>6</td>
<td>Bishen Patel</td>
<td>6</td>
<td>Software Development</td>
<td>LGY</td>
</tr>
</tbody>
</table>

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.

**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**.

![Custom AutoFilter window](image)

**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:

![Filtered spreadsheet](image)

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](http://www.ictlounge.com/html/dataanalysis.htm)

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