**Entering data**

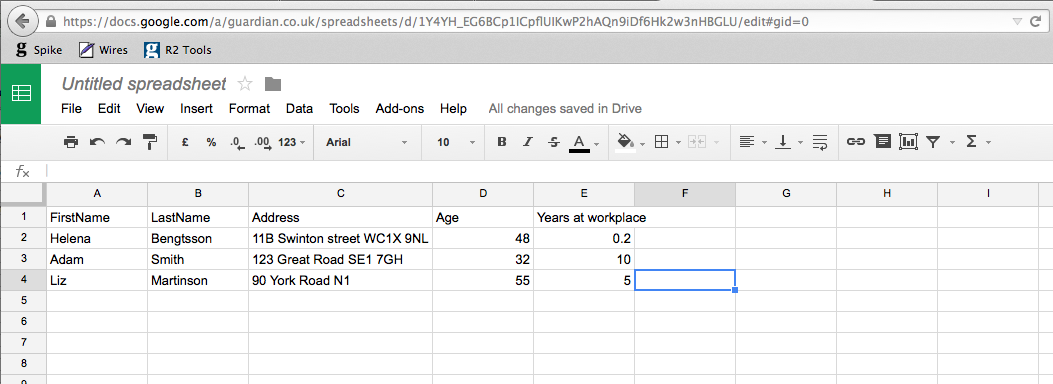
Make sure that you figure out a structure before entering the data you want in your table. Which are the important facts? Which data do I need to have in my table, and which of them can I choose not to include. What is the logical order of the information? How do I want to use the table once it’s finished? It’s of course possible to change if you should discover that you’ve made something wrong, but I advise you to do your thinking before creating your table.

As the program starts you will already get an empty sheet to work with. Start by entering the headers of your columns. Then select the header row by clicking on the number to the left of them, i.e 1, and choose Bold on the buttons in the tools field.

Include as much information as you can. You have to consider the price of your own time, but know that patterns may emerge in a totally different place than the one you thought when you started.

Put repeating data in rows - and put their properties in columns. For example: Every person is a row and the columns have properties such as first name, last name, address, income, age. A tip is to always separate first and last name. It will be much easier to sort by last name if you make one column for the first and one column for the last name.

*Exercise: Entering data. Make a simple table by entering information about the participants in the work shop. Start by opening the program and place the cursor in cell A1. enter the following headers in cells A1 through H1: Last name, First name, City, Work place, Age and Work years. Then enter all the information for each participant in the work shop in the corresponding cells.*



Enter the data you want and use the tab key to move to the next cell, cell just to the right of you. If you make an error you can change by either double clicking in the cell and changing the text in the cell or by changing the text on the function row just above your sheet. It’s often easier to make the changes there since there is more room and you don’t have to double click to change anything.

When you get to the last column you can just use the enter key or Ctrl and left arrow to get back to the beginning of the row.

The program will help you to enter data into the table. If you are making a list of people and several of them live in Stockholm you just have to type Stockholm once. When you then start typing Stockholm the next time you will notice that the program automatically fills in the rest of the word. This works whenever you have repeating information. You can also copy data from cell to cell.

### Moving around

The easiest way to move from one cell to the next is to use the arrows. But there are a couple of short keys that are good to know:

Tab Moves to the next cell to the right

Ctrl (or ⌘) + right arrow Move to the column furthest to the right on the same row

Ctrl (or ⌘) + left arrow Move back to the first column on the same row

Ctrl (or ⌘) + down arrow Move to the last row (with data) in the worksheet

Ctrl (or ⌘) + up arrow Move back to the first row (with data)

Ctrl (or ⌘) + End Move to the last cell (last column, last row) with data

Ctrl (or ⌘) + Home Move back to the first cell (first column, first row) with data

### Column width

There are several ways to change the width of a column. Move the cursor at the top of the sheet, just between two columns and you can see the cursor changing from a plus sign to a double arrow. Select and change the width of the left column. It is also possible to double click while the cursor looks like a double arrow. That makes the left column adjust it’s width to the widest cell in the column. By selecting the whole sheet (click on the small square between row 1 and column A) and then double clicking between two of the columns you make all of the columns adjusting their width on the same time.

### Freezing rows and columns

While working with a large worksheet it can be useful to be able to see your headers when you are entering data at the end of the worksheet. It is then possible to lock the headers by selecting cell A2 and then clicking on the View tab and Freeze rows. You can then freeze the top row or the top three rows if you want to. By selecting Freeze column you can freeze columns in the same way.

### Selecting

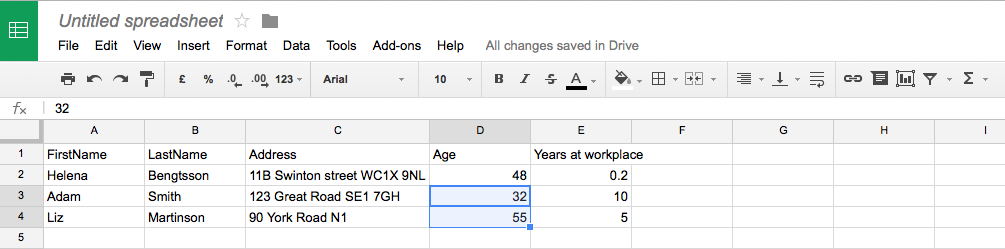
You must select a cell, row, column or sheet before doing something with it, like changing the font, the format or sorting your data. Select a cell by moving to it – or by using the mouse and clicking. Note that the name of the cell is shown just beside the function field, if you are in the third column and on the fourth row the name of the cell will be C4. You can also select an entire row by placing your mouse on the beginning of the row, at one of the numbers, and click. In the same way you can select a column by clicking on the letter above the column. Finally it is possible to select the whole worksheet by clicking in the square between row 1 and column A.

### Adding rows and columns

You can add a row by selecting the row below where you want to add the new row. Then chose Insert and choose insert row below (or above). For columns you select a column and then choose Insert and Column left (or right). Or right click on a column or row you’ve selected and then choose Insert.

### Copying

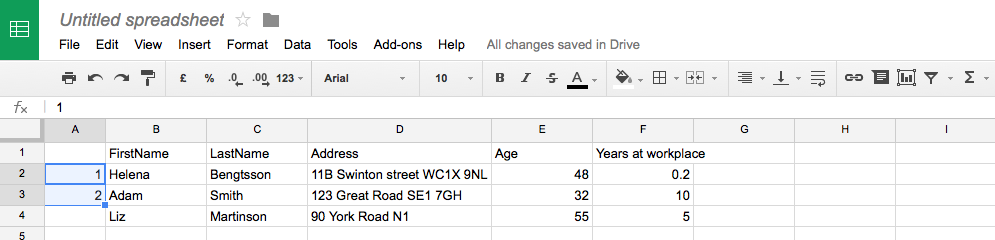
The usual short commands, Ctrl (or ⌘) -C for copying, Ctrl (or ⌘) -V for pasting, are used here too. You can also find the Copy, Cut (Ctrl (or ⌘) –X) and Paste commands in the menu at the Home tab. There is also has another way: Select a cell by clicking. Move the cursor to the lower right hand corner of the cell, the cursor changes from a arrow to a thin plus sign. You can also see a small handle in that corner, select the corner with the thin plus sign and pull down. The contents of the cell are copied to the cells next to the original cell.



### Numbering

You can use this handle while numbering the rows. Enter 1 in the first cell and 2 in the cell below. Select both cells and select the handle with the slim plus sign and pull. The program then continues the range of numbers you have begun. You can start your number series anywhere; try writing 101 in the first cell and 102 in the next. Or try 10 in the first and 20 in the second. This works with anything that has a sequence, Monday – Tuesday, January – February and so on. And with dates – this is an easy way to make a calendar – The program increments the dates with the interval you have chosen in the first two, every day, every week or every tenth day…

*Exercise: Number your table. Create a new column to the far right by placing the cursor at column A and choosing Insert on the Home tab. Add the header No. Enter the number 1 in cell A2 and the number 2 in cell A3. Select both cells and take hold of the copy handle in the second cell. Pull down and watch how the program numbers your rows from 1,2,3,4,5 ...*

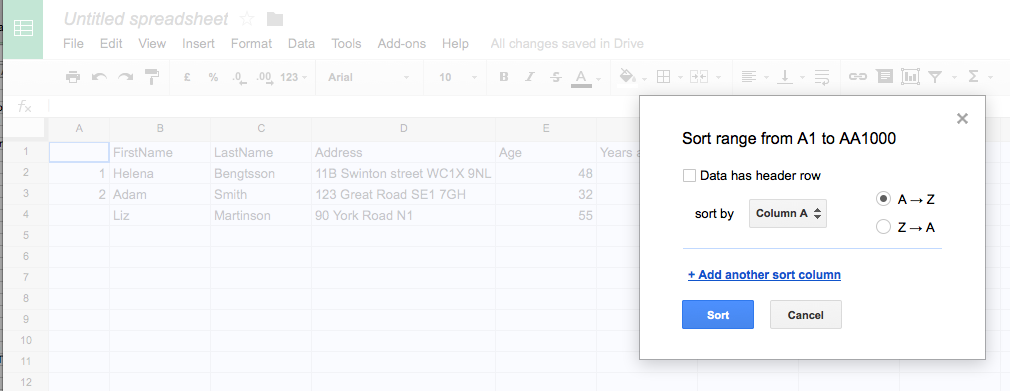


### Sorting

Start by selecting the area you want to sort. In most cases you’ll be ok with selecting the whole sheet, select that by clicking on the small square between row 1 and column A. Always include all your columns when you want to sort, otherwise there will be a risk of destroying your data. Select the whole sheet or select the rows you want to sort.

Chose Data and Sort range in the menu. A sorting window opens where you can select which column you want to sort on and also if you want to sort the column ascending or descending. (Tip: Text is often sorted ascending, A-Z, and numbers descending, biggest – smallest) If you have headers in your worksheet the program might notice that by itself, otherwise tick the box that says Data has Header Row.

*Exercise: Sort the table to find out who’s the oldest. Select all by clicking between A and 1 in the upper left corner of your sheet. Then choose Data and Sort range. Don’t forget to check Data has header row. Then use the drop down menu to chose Age and Largest to smallest to find out who’s the oldest.*



**Summing and calculations**

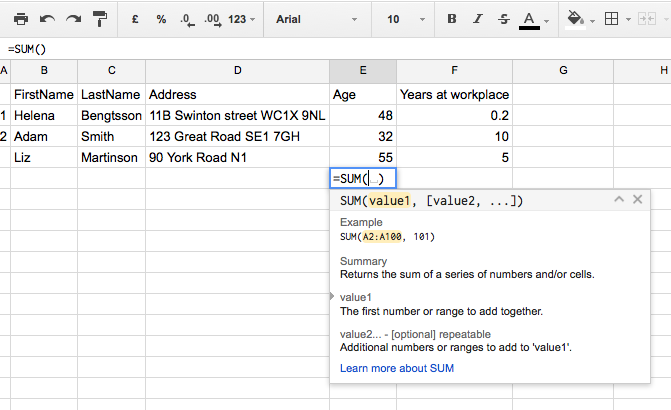
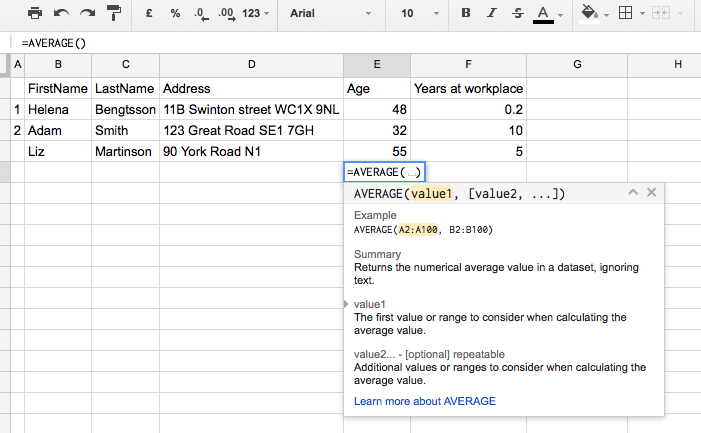
There is a special button for summing – looks like a sigma sign (AutoSum) and located on the Home bar in the left corner. Select a cell below or to the right of the cells you want to sum, click on the button and Excel suggests an area of cells to sum. You can accept that area by just pressing Enter or change the area. The easiest way to change an area is to drag the cursor over all the cells you wish to sum. It’s also possible to enter the area by changing the formula. The formula for summing is =SUM(cell to start : cell to finish), =SUM(B2:B14). This text, formula, is found in the command field just above the work sheet.

By changing SUM to AVERAGE you make the program calculate the average instead. Make sure you don’t delete the equal sign or any of the parentheses while changing the formula.

Simple calculations also start with an equal sign, and you can use numbers or cells – or a mix of the two. Example =234+12, =A3\*10, =(F5-G5)/F5

*Exercise: Calculate the collected work years of the participants. Position you cursor in the cell just below the last participant’s work years. Click on the AutoSum button and watch as the program suggest the function: =SUM(G2:G7) Press Enter to accept the function.*

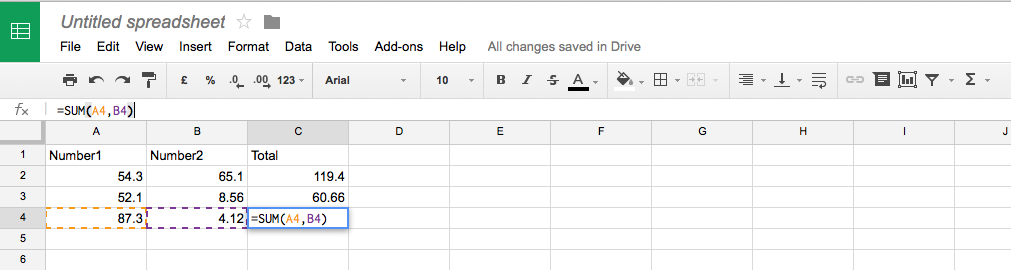
*You can also calculate the sum of peoples ages, but in that case it could be better to calculate the average age. You can use the AutoSum button here too: Place your cursor just below the ages and click on* ***Σ.*** *Press Enter to accept and then use the function field to change the function to =AVERAGE (F2:F7). Press Enter once again and you have successfully calculated the average age.*

**Copying formulas**

On of the best features with spreadsheets are their ability to change the values of a formula when the formula is copied. Say that you have several columns of numbers and you want to sum each row at the end of that. The formula for summing the first row would look something like this: =SUM(B2:F2) and would be placed in cell G2. Copy the cell with the formula to cell G3 – either by selecting cell G2, pressing Ctrl-C, moving the cursor to cell G3 and pressing Ctrl-V – or by selecting cell G2, taking hold of the small handle in the bottom right corner and dragging down. When you look at the formula in cell G3, you see that the program has changed the cells included in the formula to =SUM(B3:F3). The formula will continue to change when you continue to copy.

Sometimes you would like to keep one or both values while moving or copying a formula – the way to do that is to add a $-sign in the formula: =SUM(A$1$:A$34$)

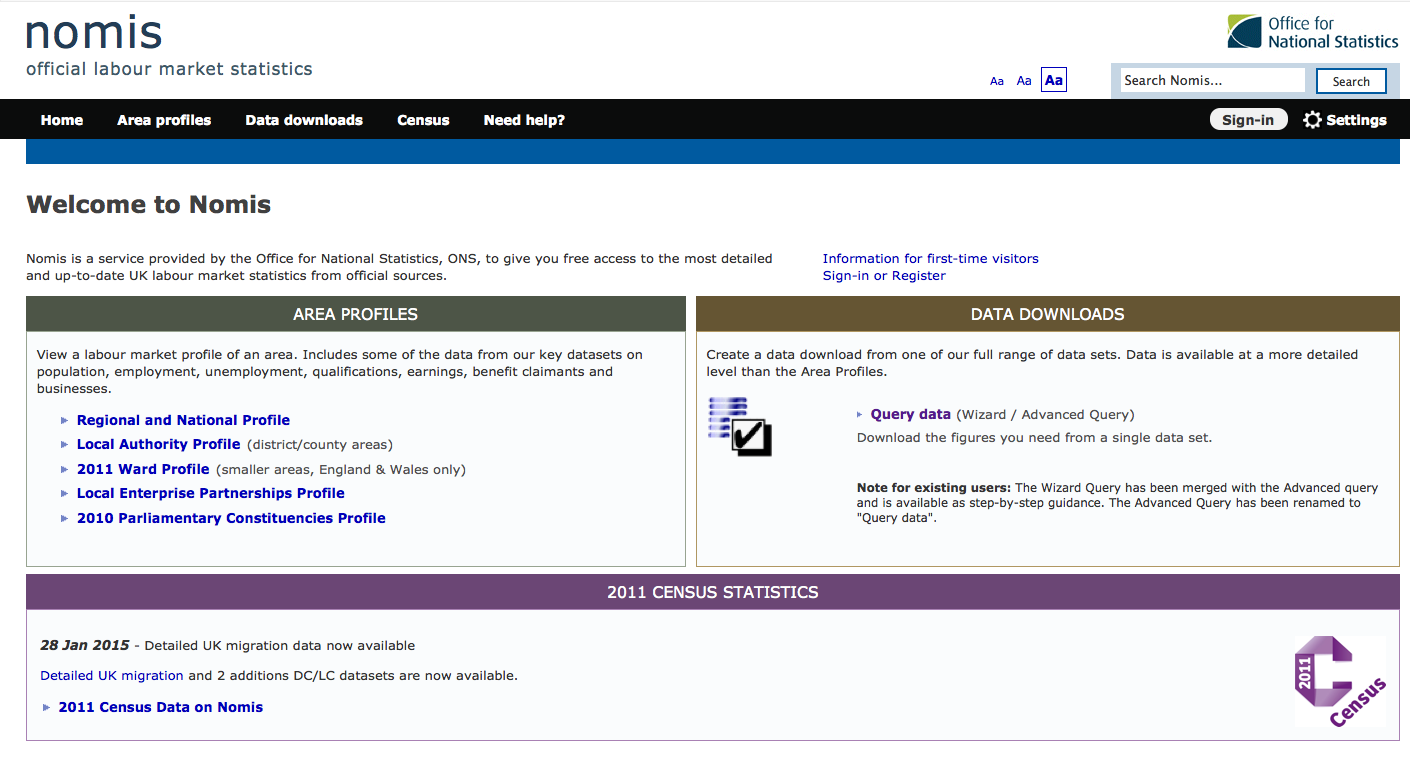


Don’t forget to save your file – even during your work. And a tip is to save several versions of the fie, sometimes you realize you’ve made a mistake when you have worked a while and then it’s great if you have a previous version saved. You save by selecting the strange looking button way up in the left corner. Use Save As to save multiple versions. Remember that Ctrl (or ⌘) - Z undoes your last key strokes.

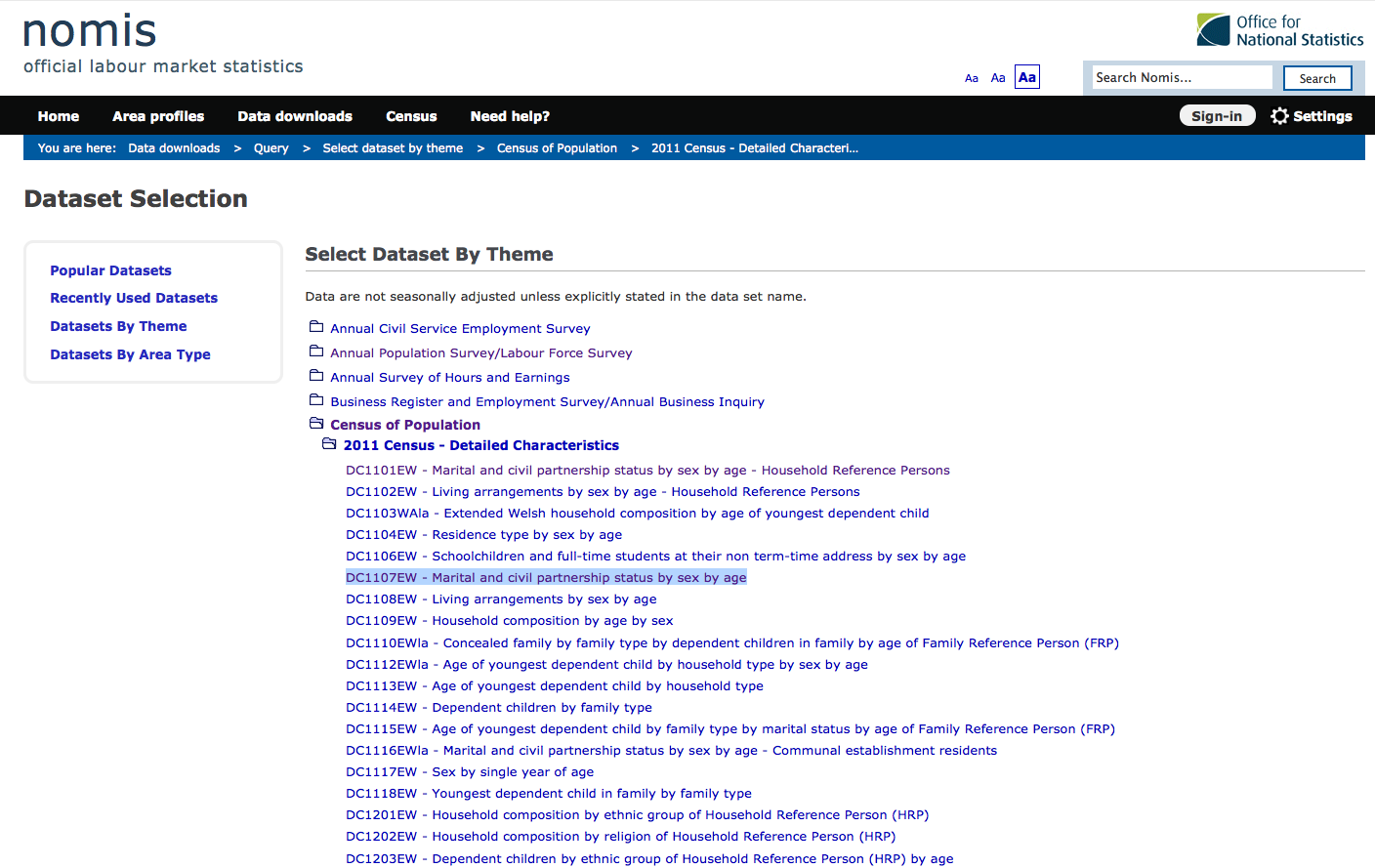
**Downloading data from the web**

In Sweden we often talk about the myth of all the women leaving the northern parts of Sweden and moving to the big cities. Wouldn’t it be fun to check out if that is true also for England and Wales? ONS has tables and databases of a lot of things, but I find that Nomis is much easier to use. There you can download parts of their databases and use for your own calculations.

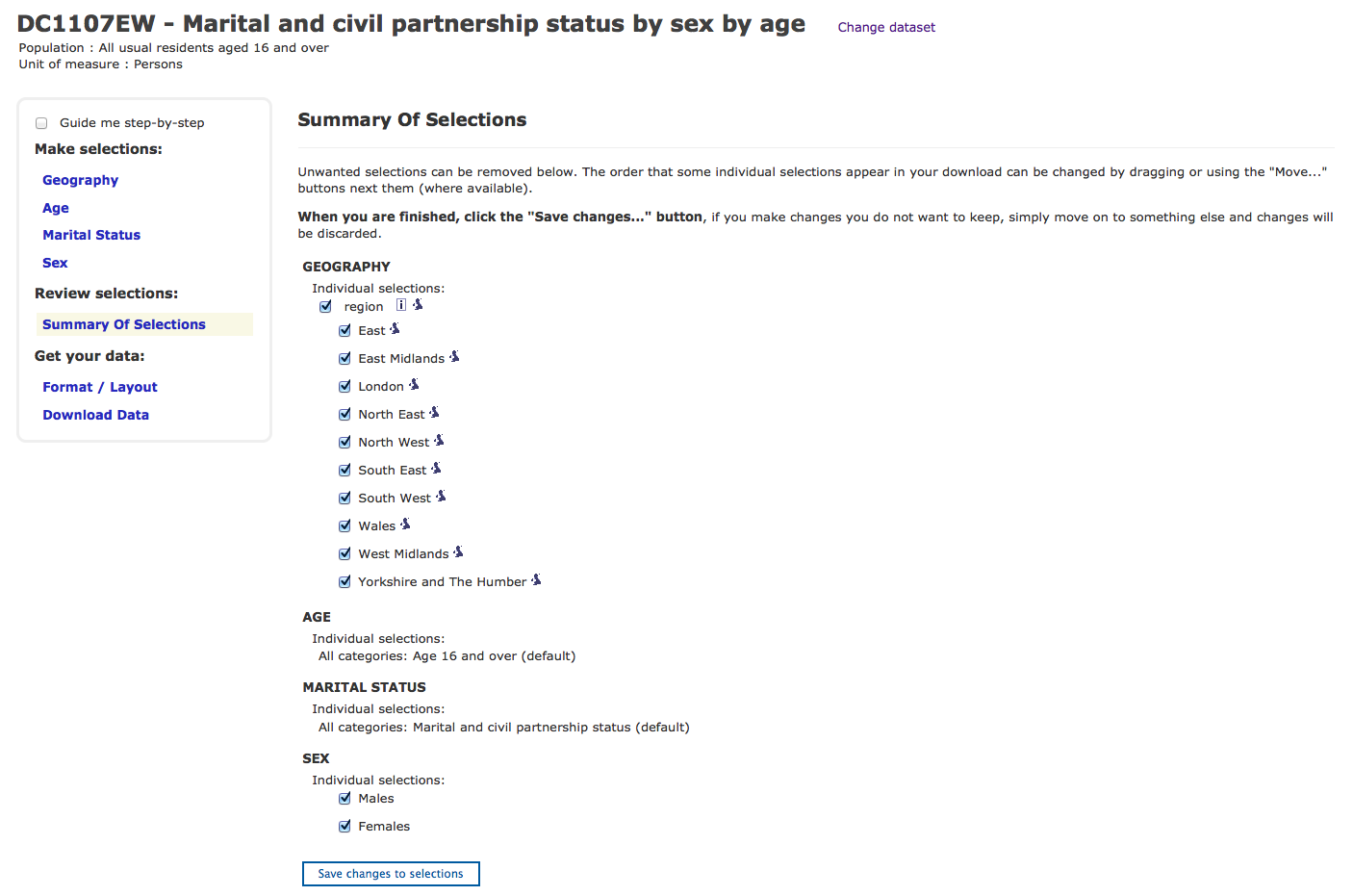
Start with opening [www.nomisweb.co.uk](http://www.nomisweb.co.uk) and choose **Query data** to the right of the page.



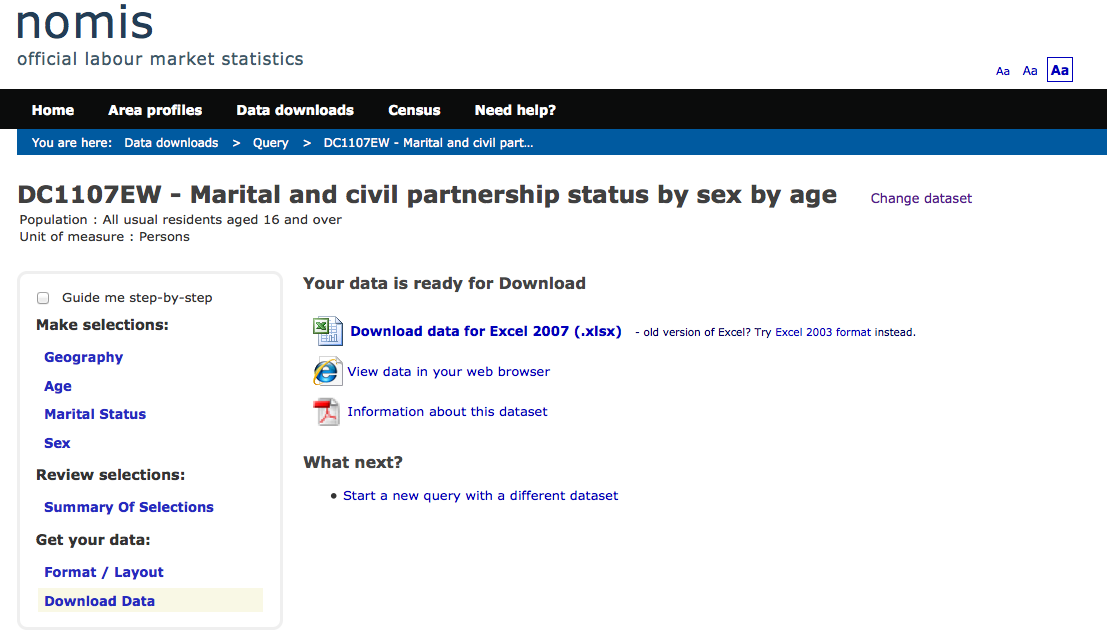
Click on **Census of Population** to choose tables containing population data. Then click on **2011 Census – Detailed Characteristics** and finally choose table **DC1107EW - Marital and civil partnership status by sex by age**. It’s now possible to select your variables:



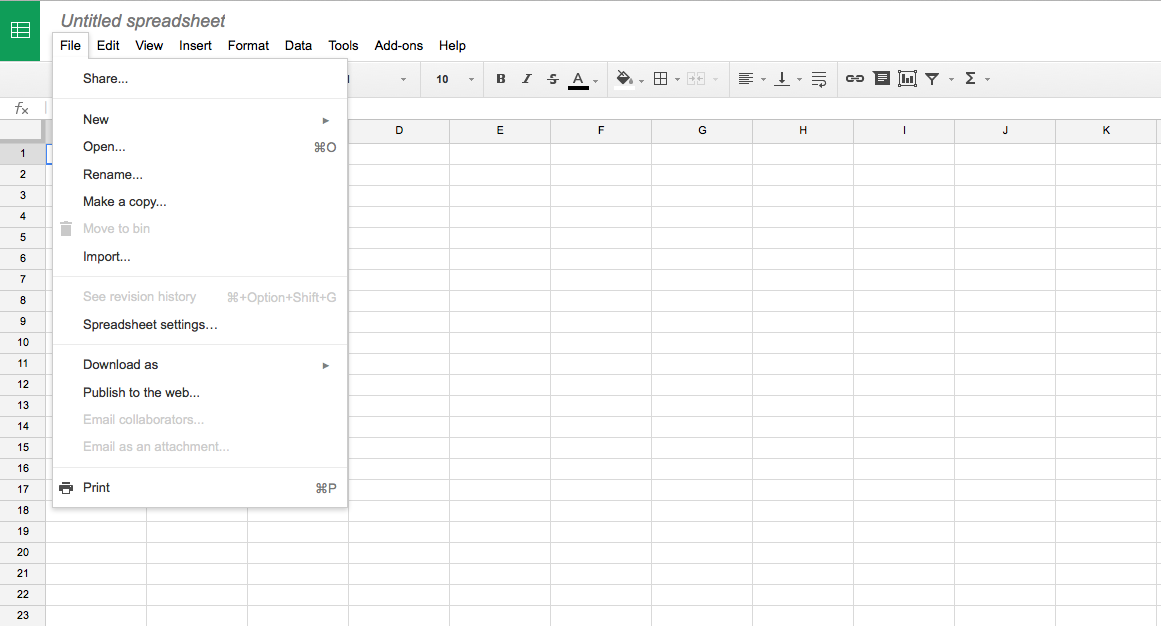
Lets start with the geography. Click on **Geography** in the box to the left of your screen and then choose the bottom one, **regions**. Click on **Tick all** and that means that you have chose all ten regions. Let’s skip Age and Marital Status for now and go straight to Sex. Click on that in the left hand box and start by un-ticking the box that says **All persons**. Then tick **Females** and **Males**. By doing that you will get one number for women and one for men, but no total. Click on **Summary of selection** to review what you’ve selected.



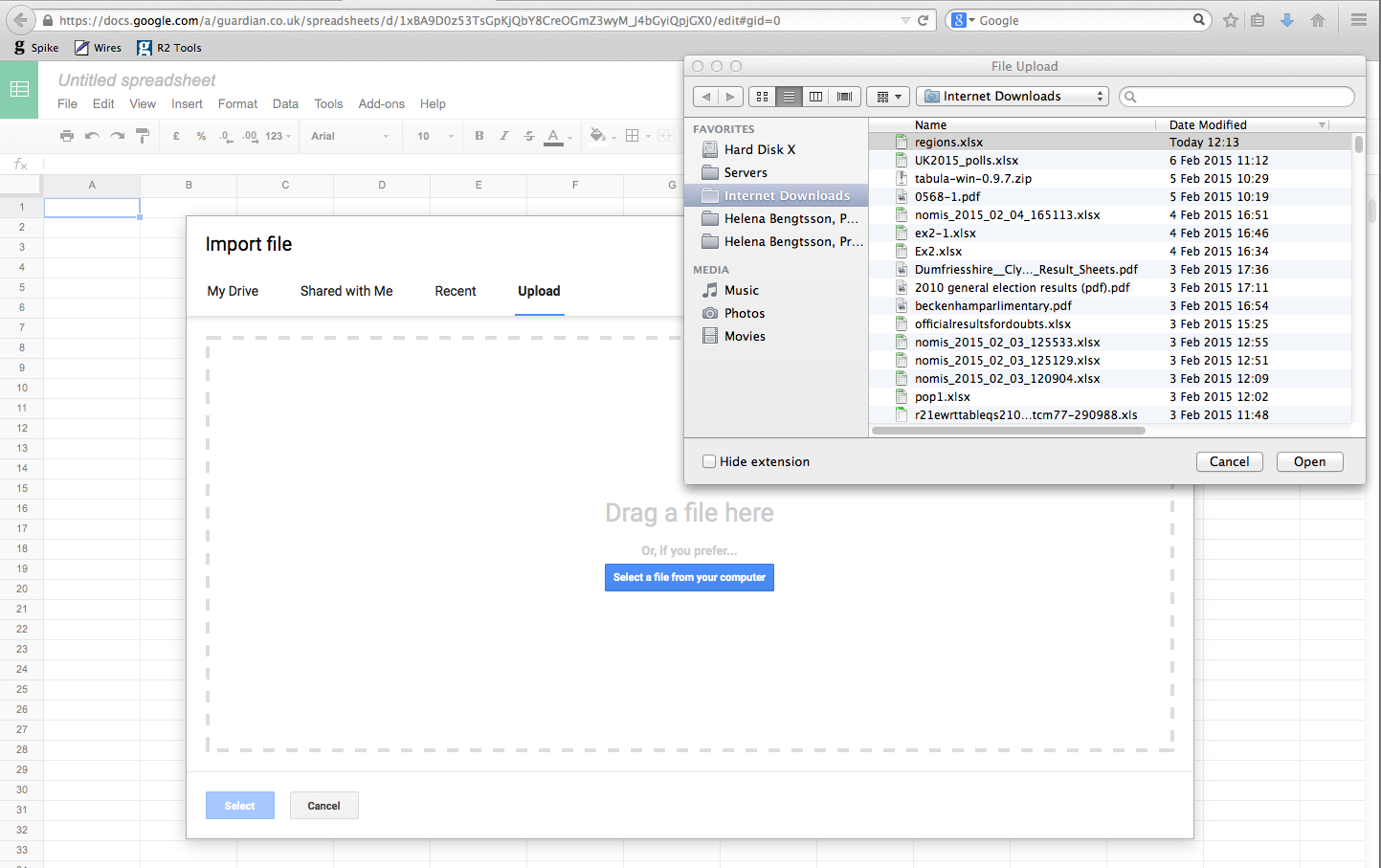
Click on **Format/Layout** and check so that there are two columns and 10 rows in your table and if you want you can also name your file on this page. Let’s give it the name regions.xls. Click on **Download data** to download.



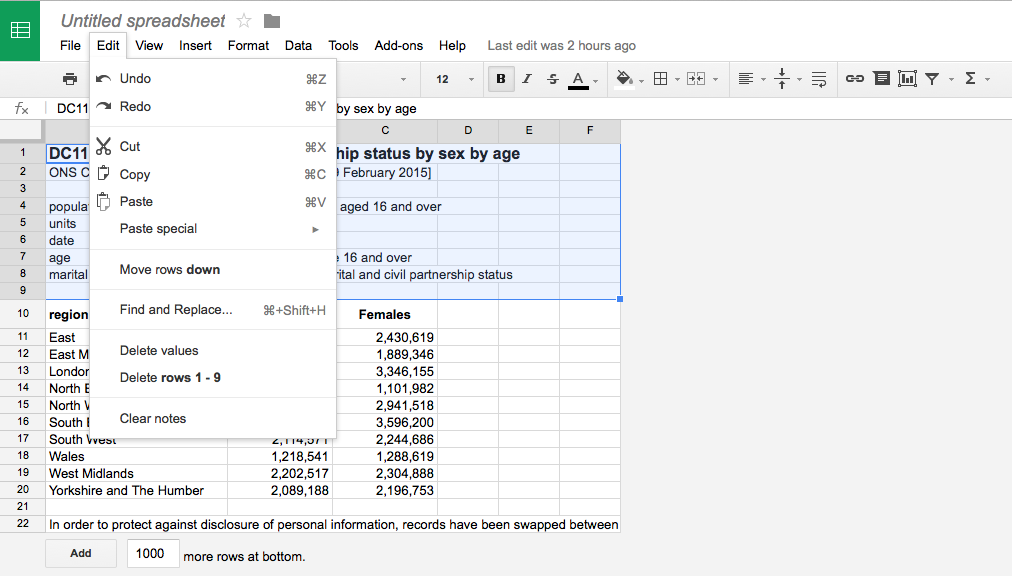
Click on the **Download data for Excel 2007** link to download and save the file. Remember where you’ve saved it. Go to Google Drive and click on the big red button that says **Create**. The arrow next to the button says Upload, but that will keep the excel-file in it’s formatting and make it harder to work with. So, even if it’s not logical, choose **Create** instead of Upload. Then choose **Spreadsheet**. Once the spreadsheet has opened click on **File** in the menu and then **Import**.



Click on **Upload** and then **Select a file from your computer**. Find the file and click Open. Choose to **Replace the speadsheet** and click **Import** to get the file into Google spreadsheets.



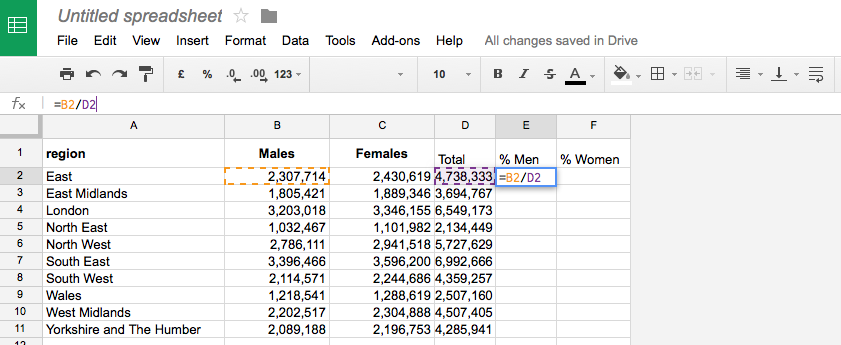
We’re going to start with cleaning the file up. Nomis and other agencies usually put a lot of information at the top and at the bottom of the file. It’s informative, but makes it hard to work with. So start by selecting the top nine rows at the top and choosing **Edit** and **Delete rows 1-9**.



Do the same thing with the last row – so that you have a clean table to work with.

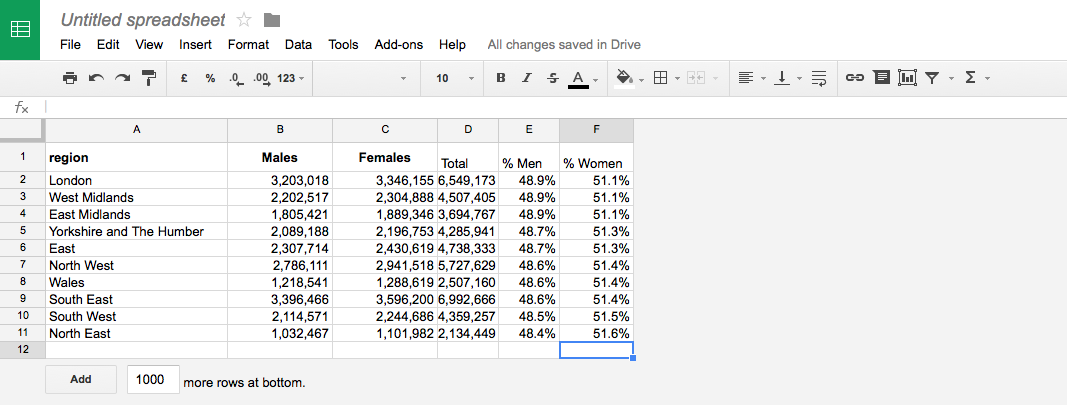
Now we can just start calculating. We want to know the percentage of the population that are women and how many that are men.

Start by calculating the total number of people in each region. Place your cursor in cell D2 and either press the autosum buttom or write **=B2+C2** in the cell. Copy your formula for all the regions by placing your cursor in cell E2 and hovering over the bottom right corner of the cell so that you see little skinny plus. Take hold of the handle and drag down to copy your function. Next we’re assigning headers for columns D, E and F. Place your cursor in cell D2 and enter **Total**. Write **% Men** in E2 and **% Women** in cell F2. Then place your cursor in cell E2, just next to the total number of people in region East. Enter **=B2/D2,** which is the number of men devided by the total number of people. Press Enter to execute your function and copy by pulling the skinny plus. Note how it changes so that row three says **=B3/D3** and **=B4/D4** for row four.



Redo the function for the women. Place you cursor in cell F2 and write **=C2/D2**. Then copy the function for all the regions. With this we have concluded the task. To clean it up a bit we could turn the entire header row to Bold so that it’s easier to read. Finally you can make the percentage numbers more readable by selecting the two columns and then clicking on the small percent sign in the middle of the menu. You can change the number of decimals with the buttons next to it.

Sort the table by first selecting the entire sheet (click between A and 1 on the top of the sheet) and then choose **Data – Sort Range** in the menu. Check **Data has header row**. Then choose to sort by % Men and choose on **Z to A** to get the highest percentage at the top.



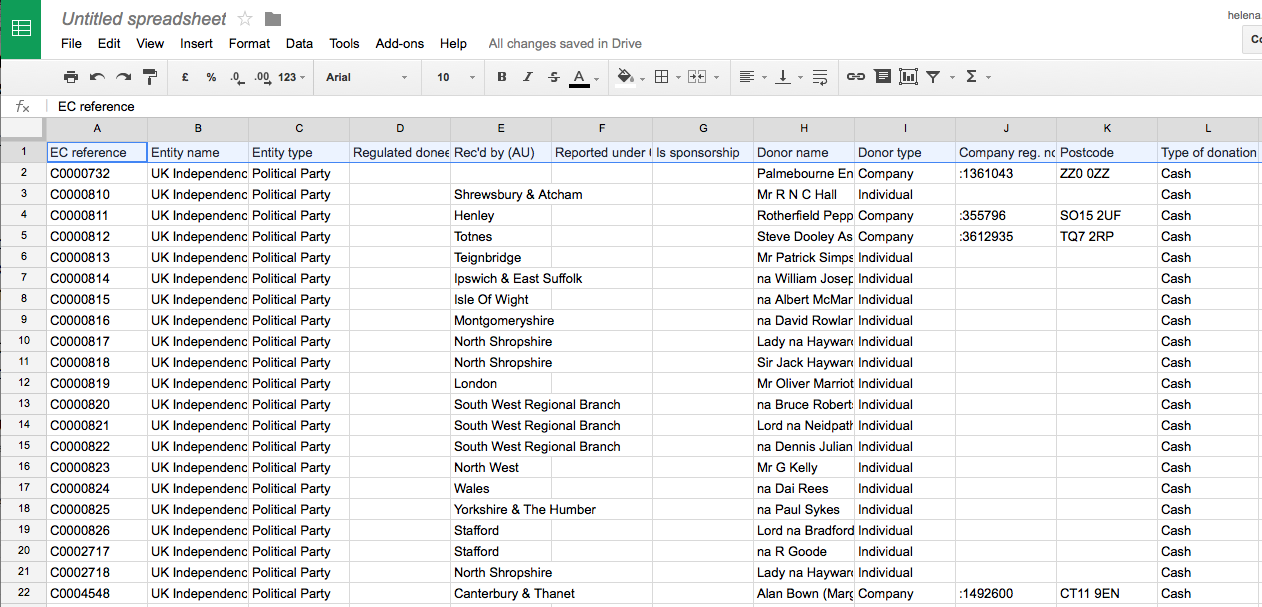
Redo the exercise by choosing all the counties instead and make it extra difficult by choosing a couple of age groups and marital statuses as well.

**Filters**

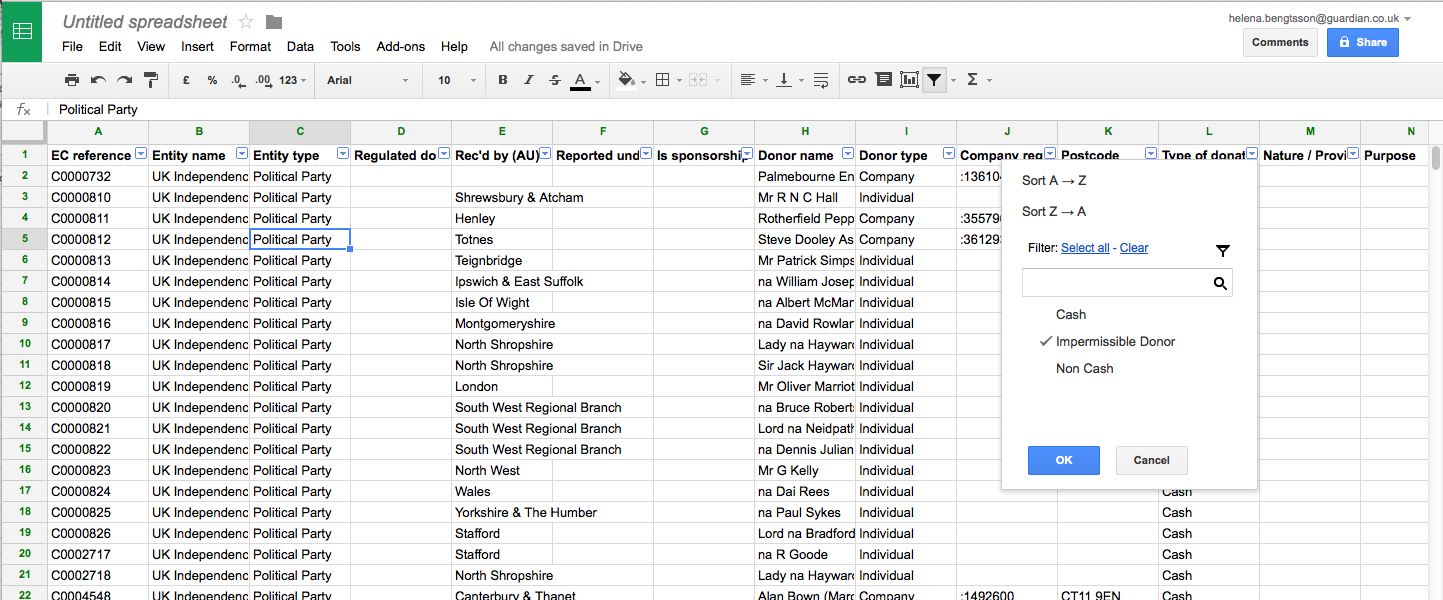
A filter is great way to “interview” your data. This is especially useful is you get a dataset from somebody and you want to get to know your data better.

Start by importing the file into Google spreadsheets by choosing Create and then Spreadsheet. Once in the new empty file, you choose File and Import. Don’t forget to click on Replace the existing sheet when you import the file.

Let’s take a look at this file:



It contains all donations given to Ukip that has been registered by the Electoral Commission. As you can see there is a lot of data there and we could use filters to take a look around the file. Start by selecting the top row and choose Data and Filter in the menu. Every headline now has a little arrow. Let’s that by selecting the arrow next to type of donation. You can then see that there are three types of donations: Cash, Non Cast and Impermissible Donor. That sounds interesting, let’s choose that. Filter lets you select a part of the file, just filtering out the rows you choose to look at. In order to do that, first click on Clear and then choose the value you want.

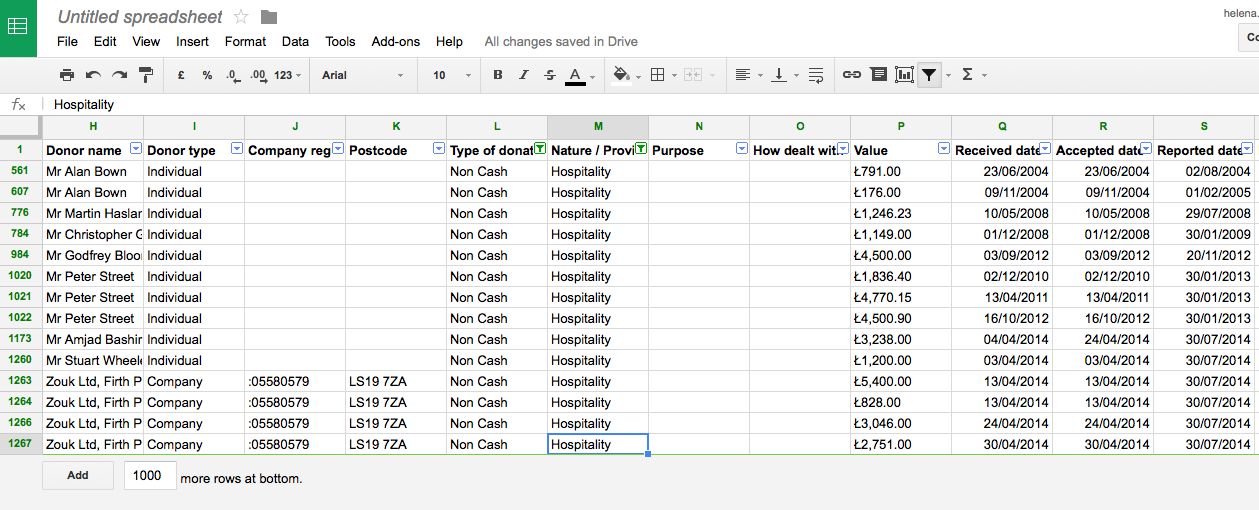


We can see that there are six rows of Impermissible donors – and if we move a little to the right in the spreadsheet we can see that all but one has been returned. Maybe there is a story there – it might be worth checking out.

Looking further though we see that the field Accepted date is empty. That is something we should look at. So click on the little arrow next to type of donation and choose Select all, and then click on the arrow next to accepted date and choose Clear and the tick next to Blanks. The same rows appear. So, we should probably omit them from our data by deleting those rows. Notice that you can’t delete all the rows when they are filtered. You will then also delete the rows between. So, delete the rows one by one by selecting the row and then choosing Edit and Delete row in the menu.

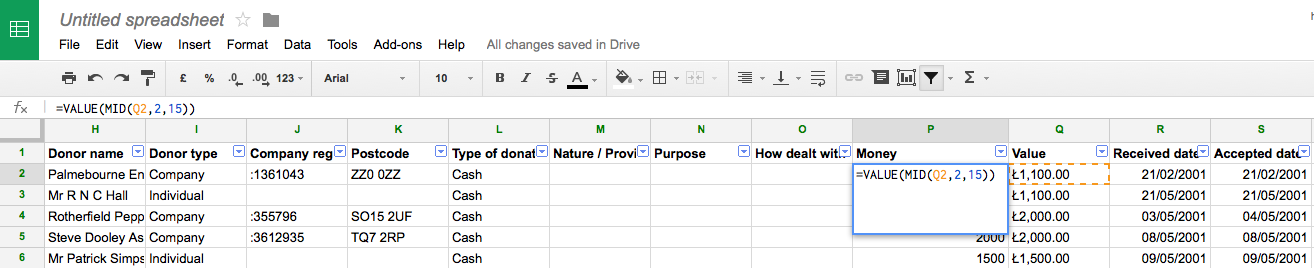
Now click on the arrow next to Accepted data and choose Select all and OK to cancel the filter.

Now let’s take a look at the Non cash donations. So go back to Type of donation and click on the arrow. Then click on Clear and tick Non cash. You see that there seem to be several different type of non cash donations under the next column Nature / Provision. Click on that arrow and you can choose among the types of non cash donations. Hospitality sounds interesting – let’s look at that. And by doing this we have filtered the list both on Type of donation and the nature of the donation. Don’t forget to roll back the filters once your done looking at your data.



It’s not possible to filter on numbers in Google spreadsheet. But you can use LibreOffice or Excel for that. It can be handy if you want to see donations over a certain amount. In Google spreadsheet you can instead solve that problem by sorting the data.

But, in this case the Electoral commission is making things a little hard for us. The Value column is actually a text column, since the sums have little strange L:s in front of them. We can create a new column and put the actual number there. Create a new column and call it Money. Then we have to write a function that does two things. It has to take away the L and it has to transform the text to a number. The function for taking something out from the middle of a text is called MID. So start by writing =MID(Q2,2,15) in cell P2. That takes the second to the fifteenth character in cell Q2 and writes that in cell P2. I choose 15 just so that I know that I will get everything. Press return and see how the result looks like. The number is still on the left which means that it’s not a number. We have to change the formula: =VALUE(MID(Q2,2,15)) will get us a number. Copy that formula all the way down the sheet.

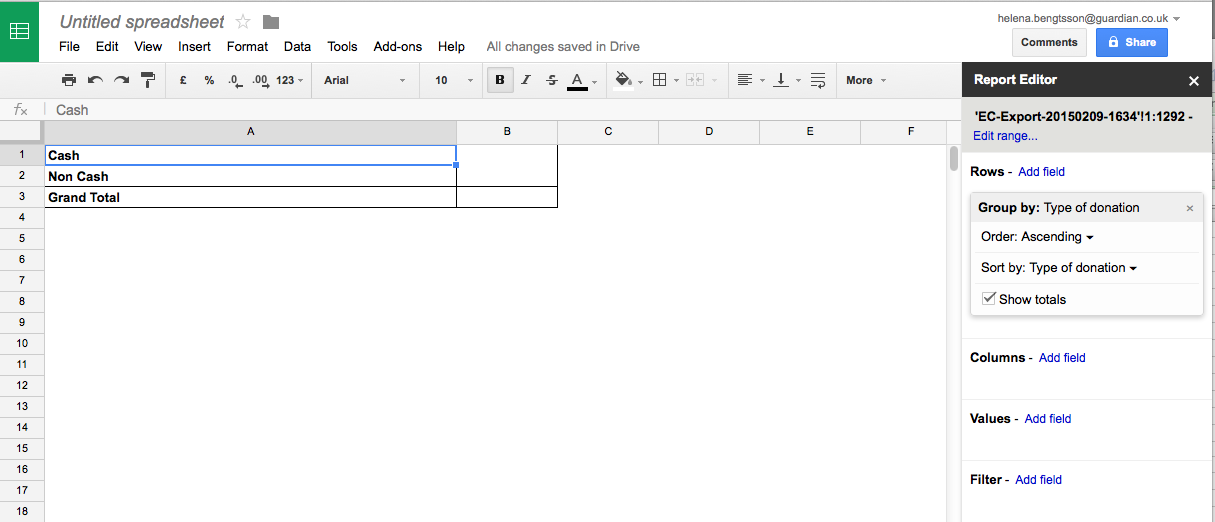


Now you can chose to sort the file after money – or by first sorting on donation type and then on money. Try looking around the file using filters and sorting and see if you can spot a story or two.

**Pivot tables**

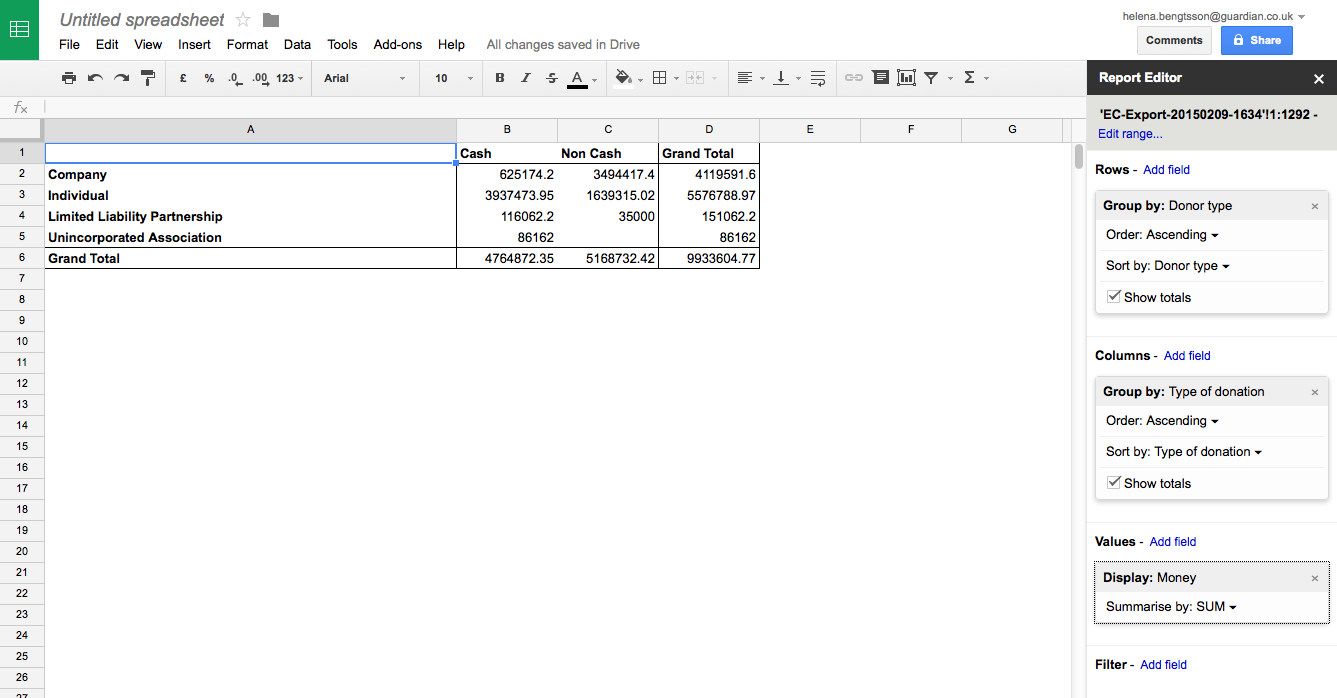
Another way to get to know your data is by doing a pivot table. Pivot tables are great for grouping and summing rows. This is how you do it.

Start by selecting your table by clicking in the square between A and 1. Then choose Tools and Pivot Table Report. You get a new sheet with two empty squares. This is where you build you pivot table. Start by adding the rows – that is the column that you want group by. For example, you would like to see how much money that was given to Ukip by cash donations, and how much that was given by non-cash donations. So let’s start there. Click on Add field next to Rows in the box to the right of the screen. You get a list of all your columns. Choose Type of donation – and you see the two values of that appear in the first empty box in your sheet.

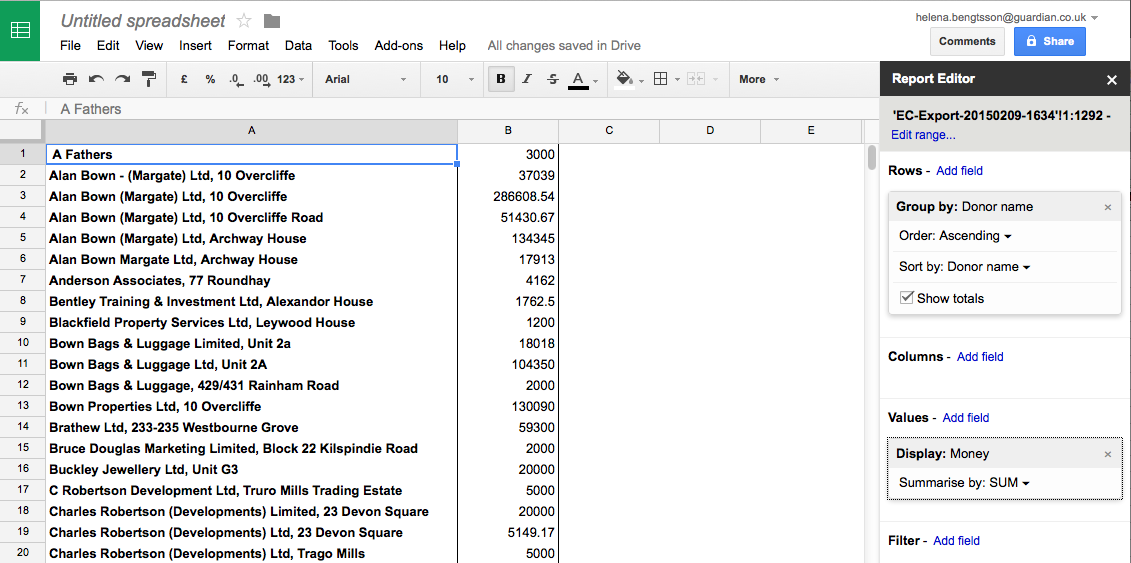


Now lets sum the money for each type of donation. Click on Add field next to Values and choose Money. And you can see the sums appear next to the types of donations and the program also gives you a total sum.

So what if we want to look at the different donor types (individuals and companies) and see weather they give cash or non-cash donations – and how much. Start by closing the boxes with types of donation and money by clicking on the small x in the corner of each box. Then click on Add field by Rows and choose Donor type. Next we want to divide the money into cash and non-cash, we do that by clicking Add field next to Columns and choosing Type of donations – and finally click on Add field next to Values and choose Money.



Now we can see how the money is distributed between both donor types and donation types. Lastly we should take a look at the individual donors. Remove all the fields by clicking on the small x:s. Then choose Donor name as your Rows and Money as your Value. And tak a look at your pivot table:



We can see that we probably will need to do a little cleaning before trying this again. It’s obvious that row 2-4 is the same person – and maybe even row 5-6 too. So, this is the fact with a lot of data – we have to clean it to get the answers. But pivot tables and filters are a good way of getting to know your data and to know what to do with it to make it answer all your questions.

In order to solve this, I would create a new column, Clean name, next to the name column. Then I would sort the sheet on the name and decide on one name for Alan Bown (Margate) Ltd and put that name next to all rows with the variations of that name. And then work my way down the list. There are tools like OpenRefine that can help you with cleaning files like this. And once you’re done you can run your pivot table once again.