

## Data Management using Excel

### Importing information from different sources

Excel can import information from many different sources - the following exercises will import information using three different techniques

- Importing from a Word document
- Importing from a simple text file
- Importing from a web page

#### Activity 1: Start Excel with a blank document

Then rename the three sheets as *carmake*, *insgroup* and *carsforsale*



### Importing from Word documents using Cut and Paste

The simplest form of import comes from the close integration of Word and Excel. A table selected in Word can be copied and then pasted into Excel. The Paste mechanism will attempt to lay the rows and columns from the Word table across cells in Excel.

#### Activity 2: Importing from Word

Start a Web browser and navigate to

<http://barryavery.com/blog/teaching/spreadsheet>

From this page choose *Session 3* and use a *right mouse click* to download the file *makemodel.doc* to your area on the network.

Open the file in Microsoft Word

Highlight both columns in Make / Model table and press the copy icon in the clipboard panel on the ribbon



#### Car producers

The following table lists information about cars – it lists the Model name of cars along with who manufactures it (the Make).

Make	Model
Peugeot	106 XN 1.1L
Renault	19 16V
Peugeot	205 1.8GRD
Volvo	240SE2.0 2l Estate
Peugeot	309 1.8 GRD
BMW	316i 1.6
BMW	318i 1.8
BMW	318i 2.0

**Activity 2 (cont.):**

Switch to the *carmake* sheet in Excel, select cell A1 and press the paste button on the clipboard panel

You should get 75 rows with

Rover Vitesse 2.7

as the last row.

Note how the Make and Model names for the rows are also pasted.

	A	B
1	Make	Model
2	Peugeot	106 XN 1.1L
3	Renault	19 16V
4	Peugeot	205 1.8GRD
5	Volvo	240SE2.0 2l Estate
6	Peugeot	309 1.8 GRD
7	BMW	316i 1.6
8	BMW	318i 1.8
9	BMW	318i SE

**Activity 3: Save your work****Importing from text files with Comma Separated Values**

For many applications a standard format for importing data is a *Comma Separated Values* file (often called CSV).

Here each row has items separated by a symbol called a *delimiter*. This can be any symbol (such as a space, quote or semicolon) – in a CSV file the delimiter is a comma.

Some files also wrap each individual data item with a quote to clarify the fields. This can be called a *qualifier*

This example uses a comma as a delimiter and a quote as a qualifier

Note how the first row provides unique names for the two columns in the file

```
"model","insgrp"
"106 XN 1.1L","5"
"19 16V","15"
"205 1.8GRD","6"
"240SE2.0 2l Estate","13"
"309 1.8 GRD","9"
"316i 1.6","15"
```

When importing CSV files, Excel offers a configurable set of import options to do with the layout of the data in the import file along with how the data should be placed across the spreadsheet cells.

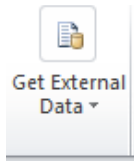
**Activity 4: Importing from a CSV file**

From the Session 3 web page use a *right mouse click* to download the file *insgrp.txt* to your area on the network.

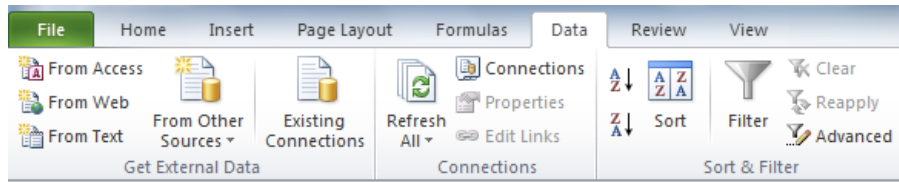
In Excel, select the *insgroup* sheet



Select the Data tab – the ribbon choices may vary depending on the size of the application window. If the page has a Get External Data icon (as below), press the arrow next to the word Data:



OR if the window is at full size, the following will be displayed



Press the *From Text* icon and find the *insgrp.txt* file using the dialog. Click the Open button once the file has been selected.

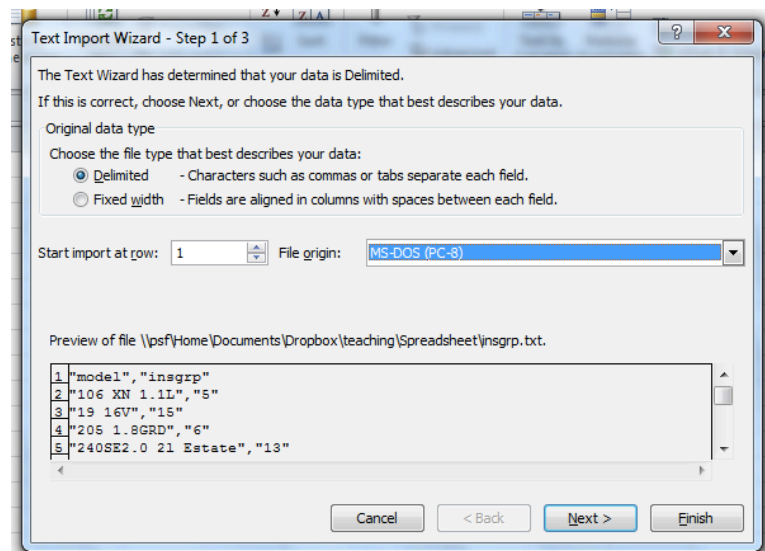
This will start the Import Data wizard

Specifying the layout of the file (Step 1 of 3 in the Wizard)

As the file uses a delimiter (i.e. each data item is separated with a comma) choose the following options

- Delimited
- Start at Row 1
- File origin (MSDOS PC-8) – This indicates that it is a plain text file with one set of items per line

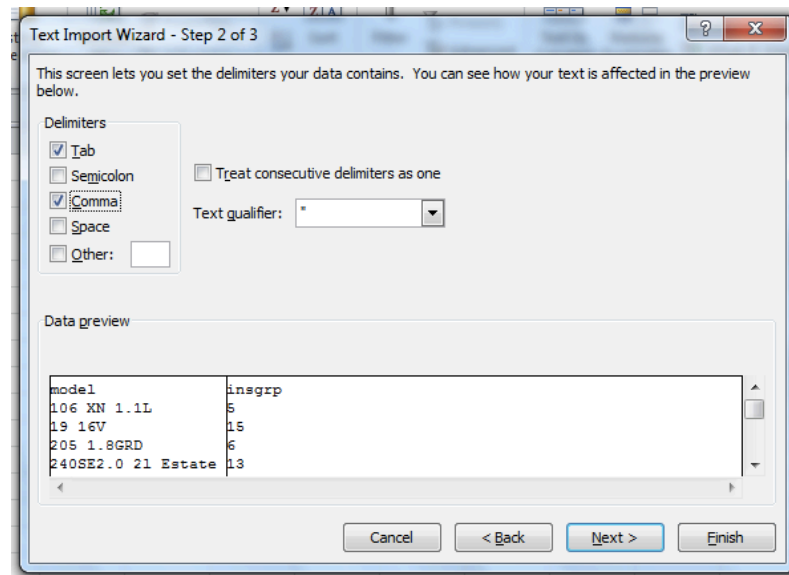
Click the *next* button



### Specifying the qualifier and delimiter (Step 2 of 3 in the Wizard)

- Choose *Comma* as the delimiter
- Choose a *double quote* as the Text qualifier (Excel may have already detected this by default)

Click the *next* button

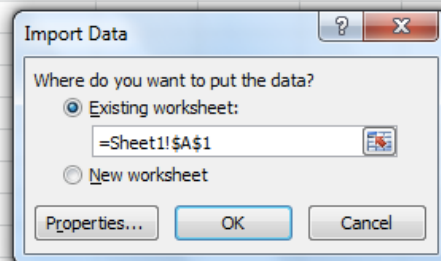


### Specifying the type of each column (Step 3 of 3 in the Wizard)

Excel will attempt to detect the type of each column in the data file – in this instance leave both columns as *General* which is the standard type of information in an Excel cell (for numbers or text)

Click the *Finish* button to (nearly !) complete the process

Excel should prompt for a location to save the incoming data – select cell A1



### Activity 5: Save your work

#### Importing from a web page directly

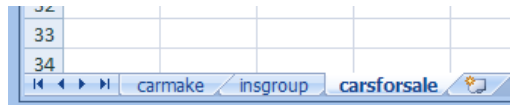
The final data source in this handout will come directly from a table in a web page.

Cars for sale - Mozilla Firefox

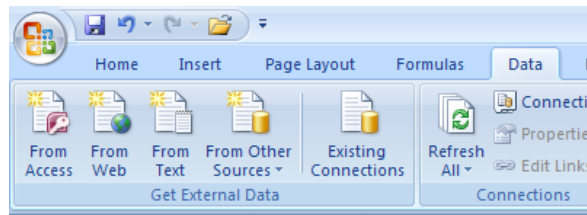
model	regno	year	price	engsize	doors	dealer	location
Rover 216 GTi	L774EWS	1993	9001	1.6	3	London West One	London W1
Escort 1.4L	K477EKJ	1993	8457	1.4	5	Curry Motors	Kingston
Astra SXi	H331CBV	1991	9775	1.8	3	Jack Brabhams	Chessington
Metro 1.1L	J123ANN	1992	5585	1.1	3	Mann Egerton	Morden

For this to work, the address of the web page which contains the data must be known in advance and the data in the page must be structured in some way so that Excel can extrapolate the cell and row information.

**Activity 6:** In Excel click on the *carsforsale* sheet



Find the Data tab as before and press the *From Web* button:



The new web query dialog box appears – this is a mini web browser that uses the Internet Explorer engine and can be used to find a web page with the required data.

In the address field type in

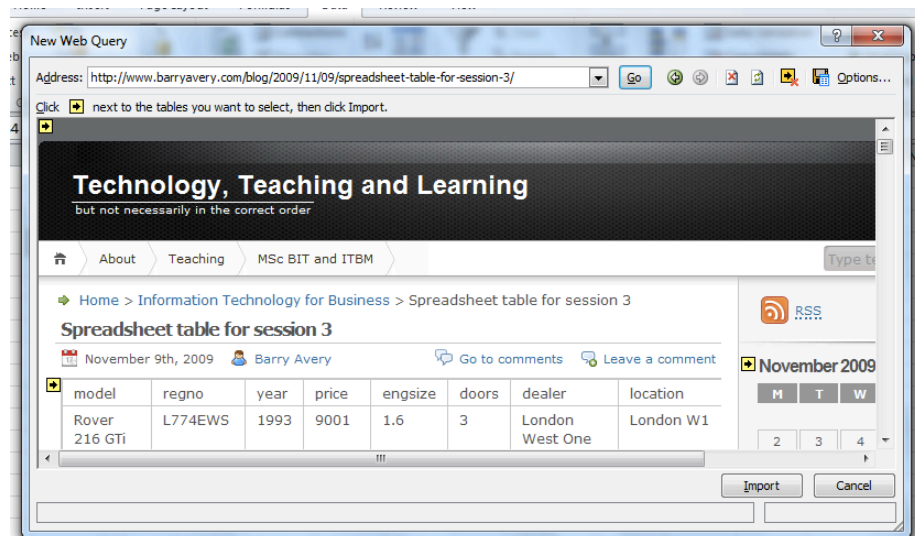
<http://www.barryavery.com/blog/2009/11/09/spreadsheet-table-for-session-3/>

This should bring up the web page at this address.

Click on the

Cars for sale

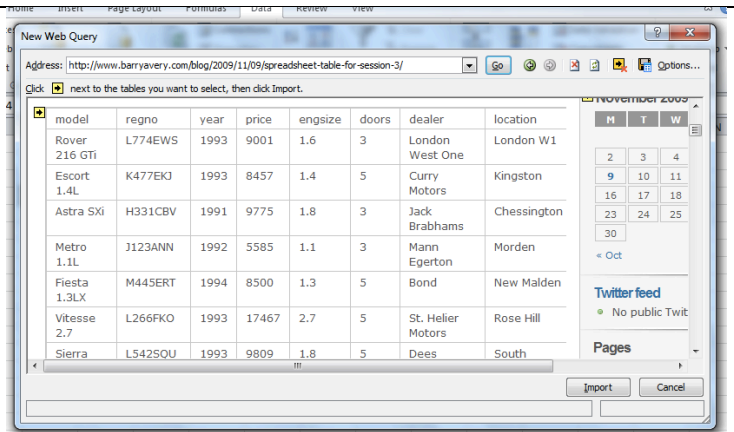
link to see the data table



Yellow arrows will appear next to the items that can be imported

Click on the arrow in the top left corner of the table to indicate that this is the area to be imported

A green tick indicates the area has been selected



Click on the *import* button and select cell A1 as the area to place the data.

**Activity 7:** Save your work.

### Merging information from different sheets

The information gathered from the 3 different sources is on three different sheets in the Spreadsheet.

- *carmake* lists the manufacturer (make) of certain models of cars
- *insgroup* lists the insurance group of certain models of cars
- *carsforsale* lists the cars on sale in a particular region

We need all this information on one sheet – probably the easiest approach is to add *manufacturer* and *insurance group* onto the *carsforsale* sheet.

**Activity 8 (Hard):** On the *carsforsale* sheet, add two columns

**OR Activity 8:** Download a merged complete version showing the linked spreadsheets from Session 3 at

<http://barryavery.com/blog/teaching/spreadsheet>

- Before the *model* column insert a column called *make*
- Between the *doors* and *dealer* column insert a column called *insgrp*

Add a *lookup* formula to bring the *make* value from the *carmake* sheet

Add a *lookup* formula to bring the *insurance group* value from the *insgroup* sheet

The look up value is *model*

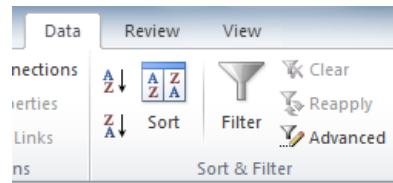
## Searching and sorting

Excel provides a number of facilities for manipulating lists. It is traditional (but not required) to have a row of unique labels at the top – called a *header row*

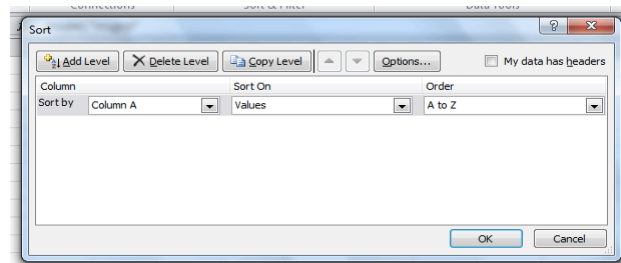
	A	B	C	D	E	F	G	H	I	J
1	make	model	regno	year	price	engsize	doors	Ins	dealer	location
2	Rover	Rover 216 GTi	L774EWS	1993	9001	1.6	3	14	London West One	London W1
3	Ford	Escort 1.4L	K477EKJ	1993	8457	1.4	5	6	Curry Motors	Kingston
4	Vauxhall	Astra SXi	H331CBV	1991	9775	1.8	3	11	Jack Brabhams	Chessington
5	Rover	Metro 1.1L	J123ANN	1992	5585	1.1	3	4	Mann Egerton	Morden
6	Ford	Fiesta 1.3LX	M445ERT	1994	8500	1.3	5	5	Bond	New Malden
7	Rover	Vitesse 2.7	L266FKO	1993	17467	2.7	5	15	St. Helier Motors	Rose Hill

## Sorting

With the cursor on any cell of the *cars* data list select the **Data** tab and click on the Sort icon in the Sort & Filter panel.



This will display the **Sort** dialog box



By selecting the fields to *Sort by*, Excel will rearrange the rows accordingly. You can sort by more than one category by using the Add Level button

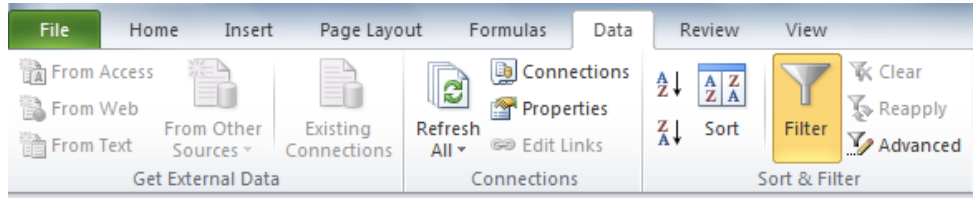
**Activity 9:** Sort the rows to list the cars by **Make** then by **Model**.

**Activity 10:** Find out how many Fords are currently on sale by Curry motors in Kingston

## Filtering

Excel provides a filtering mechanism so that only select rows that match some criteria will be displayed.

Switch filtering on by selecting the Filter icon on the Sort & Filter panel. This provides 'pull-down' buttons attached to each column label from which a selection can be made

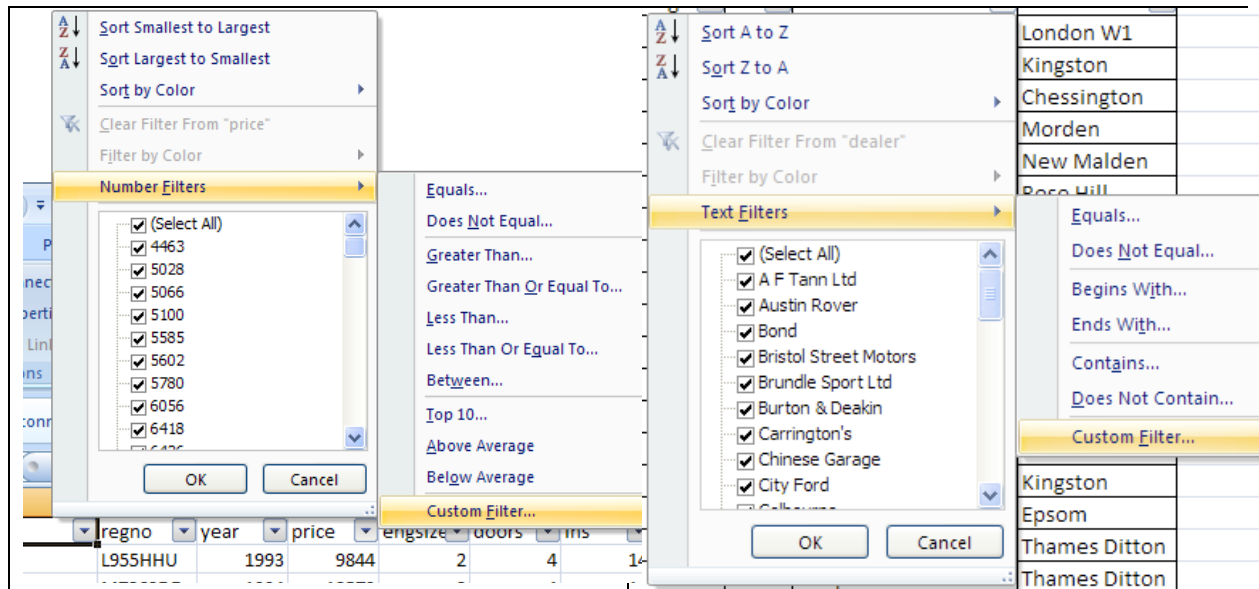


**Activity 11:** Switch on the 'autofilter' option then select *1991* in the Year column filter (this should result in 13 data rows being displayed). Now select *Ford* from the Make column to display the following table.

	A	B	C	D	E	F	G	H	I	J
1	make	model	regno	year	price	engsize	doors	ins	dealer	location
12	Ford	Granada 2.9iGhia	J83JBK	1991	6418	2.9	5	15	Bond	New Malden
30	Ford	Fiesta SX 1.1L	H895LYU	1991	8755	1.1	3	5	Kensington Ford	London
38	Ford	Fiesta 1.1L	J234ABE	1991	6796	1.1	3	5	Tony Brooks	Weybridge
39	Ford	Fiesta 1.1L	H980LAC	1991	6796	1.1	3	5	Tony Brooks	Weybridge
40	Ford	Fiesta 1.1L	H103KWE	1991	5100	1.1	3	5	Tony Brooks	Weybridge
102										

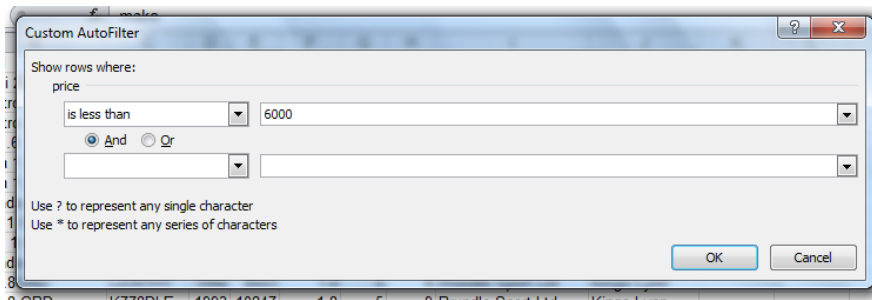
To switch the display back to all the data simply press the Filter icon.

The Custom Filter option provides access to more complex selection criteria. With Filters on, select a grey arrow and choose either Numbers Filters or Text Filters (this depends on the type of the information on the currently selected column).





To find how many cars were priced under £6,000 select -



### Activity 12: Answer the following questions

How many cars cost over £15,000?

How many Rovers cost more than £10,000?

How many 4 door Peugeot's are for sale?

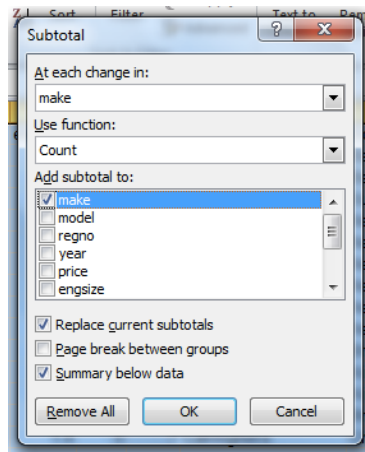
Turn off Filtering before moving onto the next section.

### Applying functions over subgroups - subtotalling

Often summary calculations need to be performed over groups of data items. Examples could be

- Show how many cars of each make are available (a count of the number of items)
- Show the total market value of cars at each garage (a sum of the cost of the cars at each garage)

Here a simple *Count* of the number of *Makes* has been performed revealing that there are 3 Audi's and 5 BMW's for sale.

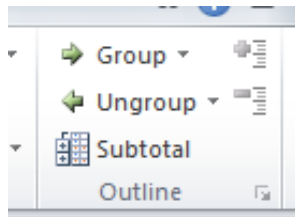


4		Audi	80TDI 1.6	L654SDF
5	<b>Audi Count</b>	3		
6		BMW	316i 1.6	M542GHR
7		BMW	318i SE	L557EFD
8		BMW	318i 1.8	M112WED
9		BMW	320i SE	J564JSE
10		BMW	320i SE	L328NEE
11	<b>BMW Count</b>	5		
12		Ford	Fiesta 1.1	K478DCV
13		Ford	Fiesta 1.3LX	M445ERT
14		Ford	Granada 2.9iGhia	J83JBK
15		Ford	Orion 1.8 LX	L569ESD

To add group functions do the following:

1. Sort the data on the required column (for example showing how many cars of each make are available would require sorting on make, showing the value at each garage would require sorting by garage).
2. Choose the functions that will be applied over the groups (such as count, sum, min, max, average etc)
3. Indicate where the function should be inserted into the sheet (typically at the change point in a group)

The Subtotal icon can be found under the Data tab, in the Outline panel.

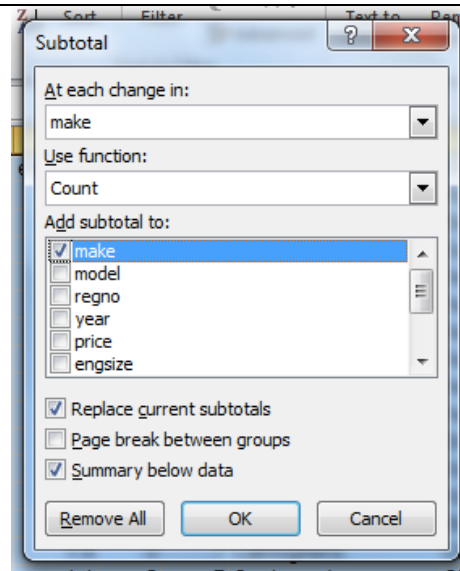


**Activity 13:** (Ensure that you have the data sorted by *make* for this exercise)

Select **Subtotals** button from the menu

Add a count after at the end of every *make*

Note: you will need to ensure that no other criteria are 'ticked' for this to work



**Activity 14:** Create subtotals for the market value of cars at each garage. You will need to sort the data according to the dealer before subtotalling

Different 'levels' of subtotals may be displayed by clicking on the 'numbered' groupings 1,2,3 found at the left of the data area (this is known as an 'Outline' feature in many packages),

**Activity 15:** Click on the outline buttons to generate the following

Select Level 1 to see a grand total

Select Level 2 to see just the subtotals

Select Level 3 to see the original data, the subtotals and the grand totals

1	2	3
1	m	
2	82	
3	M	
4		
5	M	
6		
7	Fi	
8	Fi	
9	31	
10	Or	
11	Si	
12	Gr	
13		
14	Gr	
15		
16	20	
17	30	
18		
19	21	