



**INFORMATION SYSTEMS SERVICES**  
**Data Manipulation with Microsoft Excel 2002**

**A Tutorial for Excel  
2002 for Windows**

**AUTHOR:** Information Systems Services

**DATE:** August 2004

**EDITION:** 1.0

**TUT 130**



# Contents

<b>1.0 Introduction .....</b>	<b>1</b>
1.1 About Spreadsheets .....	1
1.2 Aim of this Document.....	1
<b>2.0 Filtering Data.....</b>	<b>2</b>
2.1 Data Form .....	2
Task 1: Creating a Data Form.....	3
2.2 Autofilter.....	5
Task 2 Filtering.....	5
<b>3.0 Sorting Data.....</b>	<b>9</b>
3.1 Default sort orders .....	9
3.2 Restoring the Original Order.....	9
Task 3 Sorting Data into Groups .....	10
<b>4.0 Subtotals.....</b>	<b>12</b>
Task 4 Simple Subtotal.....	12
Task 5 Complex Subtotals .....	15
Task 6 Finishing Excel .....	18

## Format Conventions

In this document the following format conventions are used:

Commands that you must type in are shown in bold Times New Roman font.	WIN31
Menu items are given in a bold, Times New Roman font.	<b>Windows Applications</b>
Keys that you press are enclosed in angle brackets.	< <b>Enter</b> >

## Feedback

If you notice any mistakes in this document please contact the Information Officer. Email should be sent to the address [info-officer@leeds.ac.uk](mailto:info-officer@leeds.ac.uk)

## Copyright

This document is copyright University of Leeds. Permission to use material in this document should be obtained from the Information Officer (email should be sent to the address [info-officer@leeds.ac.uk](mailto:info-officer@leeds.ac.uk))

# 1.0 Introduction

## 1.1 About Spreadsheets

Spreadsheets are effectively flat-file databases, i.e. databases that contain a single table. Unlike relational databases, they cannot handle one to many relationships. For instance if information were required on a series of patient attendances at a clinic, this would be modelled in Microsoft Access by storing data about patients in one table and data about clinic visits in another.

If, however, the situation to be modelled is simple, then Excel will be a useful tool for storing and manipulating the data.

## 1.2 Aim of this Document

This document gives a set of exercises in manipulating data using Excel, and these exercises should be appropriate to anyone wanting to record and summarise data.

It is assumed that you already know how to open, close, save and print an Excel spreadsheet and that you know how to enter and edit data. If you are new to Excel it will be necessary for you to read and work through the following document before attempting these exercises:

Getting Started with Microsoft Excel 2002                      BEG 40

If you wish to investigate some of the other facilities in Excel then the following documents are also available:

Formatting with Microsoft Excel 2002                      TUT 48

Writing Formulae with Microsoft Excel 2002                TUT 47

Using Charts with Microsoft Excel 2002                    TUT 49

Before you begin, download the zip file provided for use with this document.

Open a web browser such as Internet Explorer and go to the URL:

<http://www.leeds.ac.uk/iss/documentation/tut/examples/index.html>

Scroll down to tutDataManip and right-click on the file tut130.xls. Select **Save Link As...** (**Save Target As...** in Internet Explorer). When the **Save As** dialog box appears, select a suitable directory and click **Save**.

## 2.0 Filtering Data

Filters are used to select only records (rows) that meet specific criteria. (In database terminology, filters are known as queries.) There are three different tools for finding and filtering records in Excel:

- Data Form
- Autofilter
- Advanced Filter

In this document the first two of these methods will be described. Information about Advanced Filtering can be obtained from the Excel **Help** system.

### 2.1 Data Form

A data form is a dialog box that provides a convenient way to enter, display or select a complete row of information (record). Before a data form can be used to add a record the sheet must have labels at the top of each column. Microsoft Excel uses these labels to create fields on the form. The database should be surrounded by blank cells, although it is acceptable to have the top and left edges touch the top and left edges of the worksheet.

The data form has two purposes:

- It can be used for data entry
  - New records can be added.
  - Existing records can be modified. (The changes can be undone using the **RESTORE** button as long as the record is current.)
  - The current record can be deleted. (Note, however, that once a record has been deleted, it cannot be restored and it will be necessary to re-enter the data.)
- It can be used to carry out a search for records that match the specification.

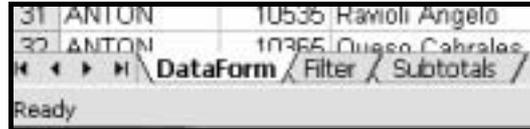
The advantage of using a data form is its ease of use. It makes entering search criteria simpler.

The disadvantages are that data forms can display a maximum of 32 fields at a time, and the complexity of criteria that can be written for record selection is quite limited.

## Task 1: Creating a Data Form

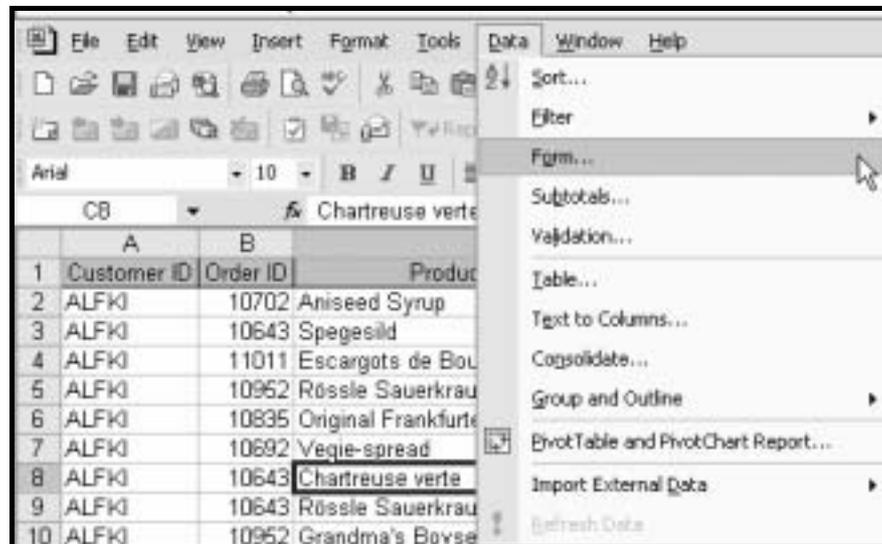
**Activity 1.1** Run Microsoft Excel by selecting **Start > Programs > Microsoft Office XP > Excel**.

**Activity 1.2** Click **File > Open > tut130.xls** (the file that you saved earlier). This file contains three worksheets. Ensure that the worksheet **DataForm** is selected.



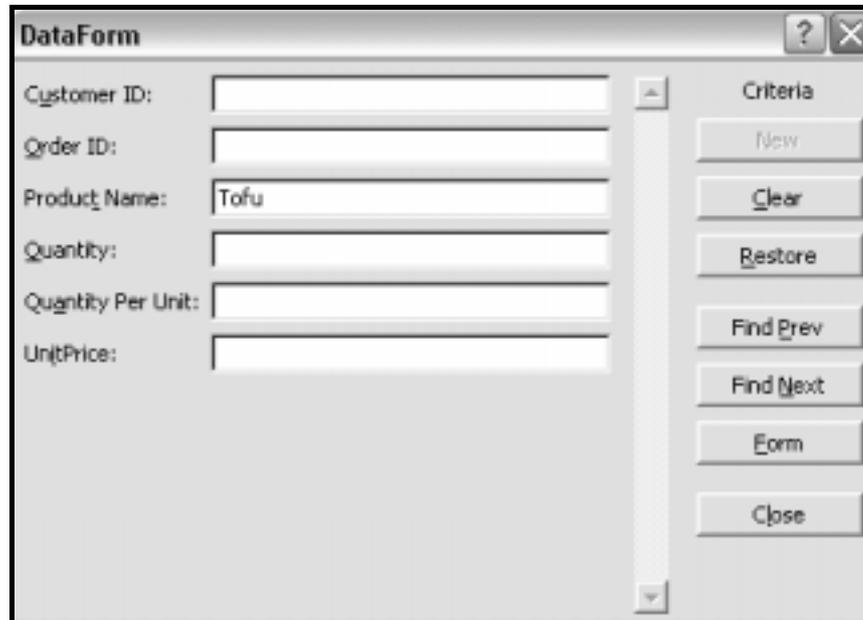
**Figure 1** Worksheet Names

**Activity 1.3** Click somewhere in the data region (A1:F2156) and select **Data > Form** from the menu.



**Figure 2** The Data Menu

**Activity 1.4** Click on the **Criteria** button and then type **Tofu** in the Product Name field. Click the **Find Next** button to move to the first record for the product Tofu. Move forwards through the matching records by repeatedly using the **Find Next** button. The last record will be for Order ID 10333



**Figure 3 Data Form Dialog Box**

**Activity 1.5** Click the **Criteria** button again to return to the earlier dialog box, and then select **Clear** to remove the word Tofu from the form.

Experiment with other searches. Note that more than one criterion can be used.

Note also that if a search is unsuccessful – i.e. no records match the search criteria, the record displayed will be the one that satisfied the previous search criteria. Care is therefore needed when using this facility.

**Activity 1.6** Use the **New** button to move to a blank record, and add new details for customer ZZZ. (If the **New** button is greyed out, click the **Form** button to make it available.) When you have completed the entry, click **New** again. Close the form and observe that your new record has appeared at the end of your worksheet.

**Activity 1.7** Select a cell somewhere inside the database and click **Data > Form** again. Use the scrollbars to move to the last record entered (Customer ZZZ). Click the **<Delete>** key to remove it.

## 2.2 Autofilter

Autofilter can be turned on simply by selecting a cell in the database and then choosing **Data > Filter > Autofilter** from the menu. (Note that if a range of cells is selected, Autofilter will filter just the selected cells rather than the whole of the database.) To remove the filters, select **Data > Filter > Autofilter** again.

Using Autofilter five choices are possible:

- **All** Show all records
- **Top 10** Display the top (or bottom) 10 records
- **Custom** Displays the Custom Filter dialog box
- **Blanks** Display only records which have a blank in that field
- **Non-blanks** Display all records with any entry in that field.

## Task 2 Filtering

**Activity 2.1** Click on the worksheet named **Filter**.

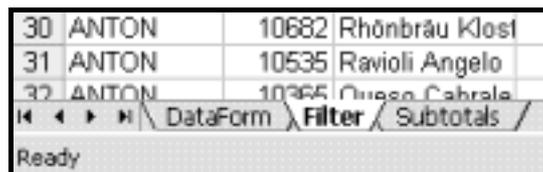


Figure 4 Worksheet Names

**Activity 2.2** Select a cell in the database and switch on **Autofilter**. The header row will now have an arrow next to each label.

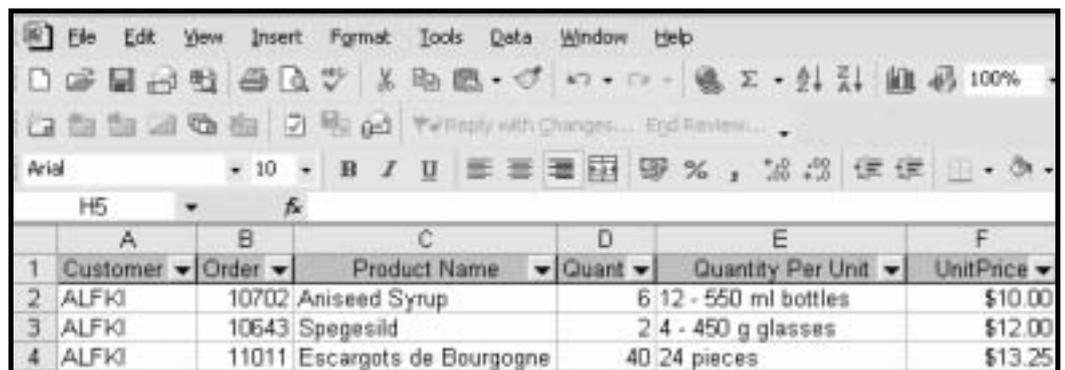
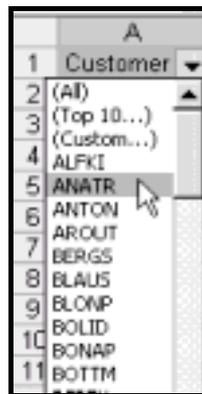


Figure 5 Autofilter On

**Activity 2.3** Click on the arrow at the top of the Customer ID field and select ANATR from the list to display only those records for that customer.



**Figure 6 Autofilter on Customer ID Field**

**Activity 2.4** Observe the result of the Autofilter. At the left of the screen, the row numbers of the selected records are displayed.

	A	B	C	D	E	F
1	Dustome	Order	Product Nai	Quant	Quantity Per U	UnitPrii
14	ANATR	10625	Tofu	3	40 - 100 g pkgs.	\$23.25
15	ANATR	10926	Teatime Choc	7	10 boxes x 12 pi	\$9.20
16	ANATR	10625	Camembert Pi	10	15 - 300 g rounds	\$34.00
17	ANATR	10926	Konbu	10	2 kg box	\$6.00
18	ANATR	10308	Gudbrandsdal	1	10 kg pkg.	\$36.00
19	ANATR	10759	Mascarpone F	10	24 - 200 g pkgs.	\$32.00
20	ANATR	10308	Outback Lage	5	24 - 355 ml bottle	\$15.00
21	ANATR	10926	Mozzarella di	10	24 - 200 g pkgs.	\$34.80
22	ANATR	10625	Singaporean H	5	32 - 1 kg pkgs.	\$14.00
23	ANATR	10926	Queso Cabrale	2	1 kg pkg.	\$21.00

**Figure 7 Filtered Records (Customer ID)**

**Activity 2.5** Now select Order 10625 to filter this subset further.

	A	B	C	D	E	F
1	Dustome	Order	Product Nai	Quant	Quantity Per U	UnitPrii
14	ANATR	(All)	Tofu	3	40 - 100 g pkgs.	\$23.25
15	ANATR	(Top 10...)	Teatime Choc	7	10 boxes x 12 pi	\$9.20
16	ANATR	(Custom...)	Camembert Pi	10	15 - 300 g rounds	\$34.00
17	ANATR	10308	Konbu	10	2 kg box	\$6.00
18	ANATR	10759	Gudbrandsdal	1	10 kg pkg.	\$36.00
19	ANATR	10926	Mascarpone F	10	24 - 200 g pkgs.	\$32.00
20	ANATR	10308	Outback Lage	5	24 - 355 ml bottle	\$15.00
21	ANATR	10926	Mozzarella di	10	24 - 200 g pkgs.	\$34.80
22	ANATR	10625	Singaporean H	5	32 - 1 kg pkgs.	\$14.00
23	ANATR	10926	Queso Cabrale	2	1 kg pkg.	\$21.00

**Figure 8 Filter on Order**

	A	B	C	D	E	F
1	Dustome	Order	Product Name	Quant	Quantity Per U	UnitPri
14	ANATR	10625	Tofu	3	40 - 100 g pkgs.	\$23.25
16	ANATR	10625	Camembert Pi	10	15 - 300 g rounds	\$34.00
22	ANATR	10625	Singaporean H	5	32 - 1 kg pkgs.	\$14.00

Figure 9 Filtered Records (Customer ID and Order)

**Activity 2.6** Now restore all the records by clicking on the Customer ID arrow and selecting **Show All**, and then by clicking on the Order ID record and selecting **Show All**. It may be necessary to use the scroll bars to move to the first option.

	A	B	C
1	Customer	Order	Product Name
14	ANATR	(All)	
16	ANATR	(Top 10...)	
22	ANATR	(Custom...)	
2157		10308	
2158		10625	
2159		10759	
2160		10926	

	A	B	C
1	Customer	Order	Product Name
14	(All)	10625	Tofu
15	(Top 10...)	10926	Teatime Choco
16	(Custom...)	10625	Camembert Pie
17	ALFKI	10926	Konbu
18	ANATR	10308	Gudbrandsdals
19	ANTON	10759	Mascarpone Fe
20	AROUT	10308	Outback Lager
21	BLONP	10926	Mozzarella di G
22	BOLID	10625	Singaporean H
23	BOTTM	10926	Queso Cabrale
2157	BSBEV		
2158	CACTU		
2159	CENTC		
2160	CHOPS		
2161	COMMI		
2162	CONSH		
2163	DRACD		

Figure 10 Show All Records

**Activity 2.7** More complex criteria can be set using the **Custom Filter** option. To select records whose Unit Price is between \$80 and \$90 inclusive, click the arrow next to Unit Price and select **Custom**. Complete the dialog box shown below:

**Custom AutoFilter**

Show rows where:

UnitPrice

is greater than or equal to 80

And  Or

is less than or equal to 90

Use ? to represent any single character  
Use \* to represent any series of characters

OK Cancel

Figure 11 Custom Autofilter Dialog Box

**Activity 2.8** It is also possible to filter text fields. To select all records for customers whose Customer ID begins with B or C, complete the Autofilter dialog box as shown. Note that the filter is not case sensitive, so there is no need to specify upper case.



**Figure 12 Custom Filter on Text Field**

**Activity 2.9** Restore all records by clicking on the arrow next to Customer ID and choosing **ALL**.

**Activity 2.10** Find a blank entry in the ProductName field by clicking on the appropriate arrow and selecting **BLANKS**. It will be necessary to scroll to the bottom of the options to find this. The following record should be returned:

	A	B	C	D	E	F
1	Customer	Order	Product Name	Quantity	Quantity Per Unit	Unit Price
	2155 WOLZA	10998		20	5 kg pkg.	\$10.00

**Figure 13 Retrieved record with Blank Cell**

**Activity 2.11** Experiment with the other options on the custom menu. Note that to be selected, **AND** selections must meet both criteria, whereas **OR** selections must meet at least one criterion.

**Activity 2.12** When you have finished, remove the filters by selecting **Data > Filter > AutoFilter**. The complete dataset will now be restored.

## 3.0 Sorting Data

There are some quirks to be aware of when using Excel's sorting facilities.

- If there is a blank line in the dataset, Excel will assume that the first blank line marks the end of the range, and it will only sort the data up to that position.
- If you have already named a range (by highlighting all the data, selecting **Insert > Name > Define** and giving the database a name) that contains one or more blank rows, when Excel sorts the data it will move any blank rows to the end.

In addition to providing sorted information for reports, it is necessary to sort a file (often on more than one field) in order to use Excel's **Subtotal** feature

### 3.1 Default sort orders

In an ascending sort, Microsoft Excel uses the following order. (In a descending sort, this sort order is reversed except for blank cells, which are always placed last.)

- **Numbers:** Numbers are sorted from the smallest negative number to the largest positive number.
- **Alphanumeric sort:** When alphanumeric text is sorted, Excel sorts from left to right, character by character. For example, if a cell contains the text "A100," Excel places the cell after a cell that contains the entry "A1" and before a cell that contains the entry "A11."

Text and text that includes numbers are sorted in the following order:

0 1 2 3 4 5 6 7 8 9 (space) ! " # \$ % & ( ) \* , . > : ; ? @ [ \ ] ^ \_ ` { | } ~ + < = > A B C  
D E F G H I J K L M N O P Q R S T U V W X Y Z

Apostrophes (') and hyphens (-) are ignored, with one exception: if two text strings are the same except for a hyphen, the text with the hyphen is sorted last.

Additional options are available: if for example a field contains months of the year, or days of the week, the data can be sorted by month or day. These lists are already known to Excel, but user-defined lists may be added if required. (Use **Tools > Options > Custom Lists** to define a new list).

### 3.2 Restoring the Original Order.

It is often desirable to return to the original sort order, after reorganizing the record sequence.

This can be carried out simply, by inserting a column at the beginning of the data, and inserting the numbers 1,2,3...

This will be demonstrated at the beginning of the next task.

## Task 3 Sorting Data into Groups

**Activity 3.1** For this task, use the sheet named **Subtotals**.

First, create a key field so that the original sort order can be restored.

Click in cell A1 and select **Insert > Column**.

Type the word **KEY** in the new cell A1.

Type **1** in A2 and **2** in A3, then press **Return** to confirm the entry.

Highlight the cells A2 and A3, and point the mouse at the bottom right-hand side of cell A3 (notice that there is a black rectangle in this position).

When the cursor shape changes to a crosshair, double click the mouse, and the series will be continued to the end of the dataset.

	A	B	C
1	Key	Customer ID	Order ID
2	1	ALFKI	10692
3	2	ANATR	10308
4	3	ANTON	10573
5	4	AROUT	10453
6	5	AROUT	10707
7	6	BERGS	10444
8	7	BERGS	10572

**Figure 14** Insertion of a Key Field

**Activity 3.2** A simple sort on a single column can be carried out using the Ascending and Descending tools on the toolbar.

Try sorting on several fields. When you have finished, sort the **KEY** column (column A) in ascending order.

Note that a sort can be undone (immediately afterwards) by selecting **Edit > Undo**.



**Figure 15** Sort Ascending /Descending Tools

**Activity 3.3** Select **Data > Sort**, and complete the dialog box as shown to sort first on Country, then on City and finally on TotalValue.



**Figure 16** Sort Dialog Box

Note that the **Options** button on this dialog box would allow sorting by month or day.

**Activity 3.4** Restore the original sort order by sorting on the Key column.



**Figure 17** Restore Original Order

## 4.0 Subtotals

Although subtotals can be created manually using the subtotal function, a facility exists to perform this process automatically. In addition to calculating a sum, there are options to calculate, for instance, a count, average, maximum value and minimum value for the specified fields.

Before using the Subtotal command, it is necessary to sort the file into the correct order.

An outline will be produced for each subtotal, which will allow the detailed information to be hidden. This technique is useful for creating charts that contain summary information.

### Task 4 Simple Subtotal

The **Subtotals** worksheet should be used for this task

**Activity 4.1** Sort the file in ascending order of Customer ID using the technique covered in **Task 3**. From the menu select **Data > Subtotals**, complete the dialog box as shown below and click **OK**.



Figure 18 The Subtotal Dialog Box

**Activity 4.2** Excel has created an outline, shown at the left-hand side of **Figure 19** below. Click on each of the minus signs to remove the detail level.

	A	B	C	D	E	F	G	H	
1	Key	Customer ID	Order ID	Product Name	Quantity	Quantity Per Unit	UnitPrice4	TotalValue	
2	1	ALFKI	10692	Vegie-spread	20	15 - 625 g jars	\$43.90	\$878.00	
3	37	ALFKI	10643	Spagesild	2	4 - 450 g glasses	\$12.00	\$24.00	
4	40	ALFKI	10702	Lakkalikötri	15	500 ml	\$18.00	\$270.00	
5	48	ALFKI	10643	Chatreuse vert	21	750 cc per bottle	\$18.00	\$378.00	
6	55	ALFKI	11011	Escargots de E	40	24 pieces	\$13.25	\$530.00	
7	56	ALFKI	10835	Raclette Courd	15	5 kg pkg.	\$55.00	\$825.00	
8	68	ALFKI	10835	Original Frankfi	2	12 boxes	\$13.00	\$26.00	
9	69	ALFKI	10952	Rossle Sauerki	2	25 - 825 g cans	\$45.60	\$91.20	
10	70	ALFKI	10643	Rossle Sauerki	15	25 - 825 g cans	\$45.60	\$684.00	
11	<b>ALFKI Total</b>								\$3,706.20
12	2	ANATR	10308	Outback Lager	5	24 - 355 ml bottles	\$15.00	\$75.00	
13	57	ANATR	10625	Camembert Pie	10	15 - 300 g rounds	\$34.00	\$340.00	
14	87	ANATR	10759	Mascarpone Fa	10	24 - 200 g pkgs.	\$32.00	\$320.00	
15	88	ANATR	10826	Mozzarella di G	10	24 - 200 g pkgs.	\$34.80	\$348.00	
16	<b>ANATR Total</b>								\$1,083.00

**Figure 19** Subtotals with Outline

**Activity 4.3** Hide columns C to H by highlighting them and selecting **Format > Columns > Hide** from the menu. A summary of the Total value of each customer's order is now displayed.

	A	B	G	H
1	Key	Customer ID	UnitPrice4	TotalValue
11		<b>ALFKI Total</b>		\$3,706.20
16		<b>ANATR Total</b>		\$1,083.00
23		<b>ANTON Total</b>		\$3,613.90
42		<b>AROUT Total</b>		\$8,272.65
73		<b>BERGS Total</b>		\$16,551.05
83		<b>BLAUS Total</b>		\$2,582.80
106		<b>BLONP Total</b>		\$20,871.25
112		<b>BOLID Total</b>		\$5,015.30
113		<b>Grand Total</b>		\$61,696.15

**Figure 20** Summary of Total Order Value for Customers

**Activity 4.4** Highlight the cells shown below. (To highlight a range of non-contiguous cells, highlight the first column, hold down the Control Key and then highlight the next column.) Click on the Chart Icon on the toolbar to activate the wizard and create a chart.

	A	B	G	H
1	Key	Customer ID	UnitPrice4	TotalValue
11		<b>ALFKI Total</b>		\$3,706.20
16		<b>ANATR Total</b>		\$1,083.00
23		<b>ANTON Total</b>		\$3,613.90
42		<b>AROUT Total</b>		\$8,272.65
73		<b>BERGS Total</b>		\$16,551.05
83		<b>BLAUS Total</b>		\$2,582.80
106		<b>BLONP Total</b>		\$20,871.25
112		<b>BOLID Total</b>		\$5,015.30

**Figure 21** Highlight the Summary Data

**Activity 4.5** Click on the minus sign corresponding to Level 1 to display only the Grand Total.

1	2	3	A	B	G	H
	1	Key	Customer ID	UnitPrice4	TotalValue	
+	113		<b>Grand Total</b>			\$61,696.15

**Figure 22 Grand Total of Order Values**

**Activity 4.6** Ensure that one of the cells in the database is selected. Remove all totals by selecting **Data > Subtotals** and then clicking **Remove All**.



**Figure 23 Subtotal Dialog Box**

**Activity 4.7** Sort the data using the field **Key** to restore the file to its original state. Highlight columns B and I and select **Format > Column > Unhide** to view the complete dataset.

## Task 5 Complex Subtotals

The next exercise will illustrate a more complex summary, using more than one field in the database. This exercise will also use the **Subtotals** worksheet.

**Activity 5.1** First sort (ascending) the data by Country and then by Contact Name using **Data > Sort**



Figure 24 Sort Dialog Box

**Activity 5.2** First use **Data > Subtotals** to create a subtotal of Total Value on the Country field, using an **average** function. Click **OK**



Figure 25 Subtotal Dialog Box

**Activity 5.3** Use **Data > Subtotals** to see the Subtotal dialog box again. This time uncheck the **Replace Current Subtotals** check box, and specify the average function again, at each change in Contact Name. Click **OK**



**Figure 26 Subtotal Dialog Box Showing Second Specification**

**Activity 5.4** Highlight columns A to G and choose **Format > Column > Hide** from the menu. Highlight columns I to K and choose **Format > Column > Hide** from the menu.

**Activity 5.5** Click on the minus signs for Level 3 to remove the details of individual transactions and view the average for each contact only. (When you do this the minus signs will change to plus signs.)



**Figure 27 Level 3 Outline Collapsed**

**Activity 5.6** Click on the minus signs for Level 2 to remove the details of individual contacts and show only the average for each contact.

1	2	3	4	H	L	M
	1	TotalValue		Country	Contact Name	
+	27	\$536.07		Australia Average		
+	32	\$64.50		Brazil Average		
+	45	\$388.13		Canada Average		
+	50	\$140.00		Denmark Average		
+	60	\$339.22		Finland Average		
+	81	\$983.68		France Average		
+	107	\$846.55		Germany Average		
+	127	\$500.68		Italy Average		
-	128	\$598.99			Grand Average	
	129	\$598.99		Grand Average		

**Figure 28** Level 2 Outline Collapsed

**Activity 5.7** Click on the minus sign for Level 1 to remove the details of individual countries and show only the overall average.

**Activity 5.8** Remove all subtotals. (Select **Data > Subtotals** and click on **Remove All**). Click in the cell above the row number and to the left of the column letter. The whole of the spreadsheet will be highlighted.

1	2	3	4	H
	1	TotalValue		
	2	\$878.00		
	3	\$75.00		

**Figure 29** Selection Cell

Click **Format > Column /Unhide** from the menu to view all the data.

	H	L
1	TotalValue	Country
2	\$878.00	Australia
3	\$75.00	Australia
4	\$702.00	Australia
5	\$375.00	Australia
6	\$420.00	Australia
7	\$390.00	Australia
8	\$209.40	Australia
9	\$351.20	Australia
10	\$625.00	Australia

**Figure 30** Columns H - L Highlighted

**Activity 5.9** Sort the file on the **KEY** field to restore the original order.

## Task 6 Finishing Excel

You should always quit any computer program when you have finished your session. Never switch off the computer without closing down Windows. Do not leave a computer whilst you are still logged on to it as others may use your account and could damage your files.

**Activity 5.1** Select the **File > Exit**. Excel will prompt you to save your document before it lets you quit. Click **<Yes>** to save the changes, **<No>** to quit without saving the changes, or **<Cancel>** to return to the spreadsheet.

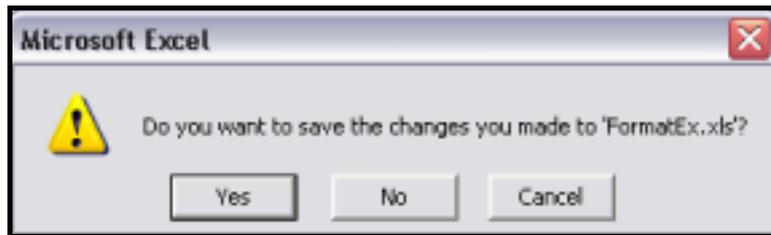


Figure 31 Quit dialog Box