Advanced Excel Skills

Note : This tutorial is based upon MSExcel 2000. If you are using MSExcel 2002, there may be some operations which look slightly different (e.g. pivot tables), but the same principles apply.

Excel Tips for Data Management and Analysis

The following document will explain some Excel functions that are extremely useful when manipulating and analysing large datasets (e.g. coverage, expenditures…). It focuses on skills that can be applied preparing for or developing Financial Sustainability Plans, as well as more broadly for managing any spreadsheet-based data.

➤ The first section will review some advanced Excel functions that simplify data management and avoid common mistakes during spreadsheet manipulations.
➤ The second section will inform you about the Analysis ToolPak that you can install on your computer for easy data and statistical analysis.

An Excel file (**Excel Advanced-English.xls**) can be used to familiarise yourself with these functions.

A – DATA MANAGMENT

Many of the data management function are located in you menu bar under DATA.
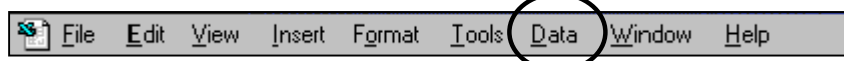
The key functions reviewed in this document will be:
➤ Sort;
➤ AutoFilters;
➤ Subtotals and,
➤ Pivot Tables.

Most of these functions can save you hours of work and minimize data manipulation errors.

A.1 – Essential Short Keys

Before learning about some of the Excel functions listed above, it is useful to review some essential Excel Short keys.

These will help you navigate through your worksheet with great ease and edit/manipulate your data much faster.

Navigation Keys - Moving faster through your Worksheet/Workbook

Ctrl + End	Moves to the last cell in the used area of the worksheet
Ctrl + Home	Moves to the upper-left corner of the worksheet
Ctrl + Page Up	Moves to the previous worksheet in the workbook
Ctrl + Page Down	Moves to the next worksheet in the workbook
Ctrl + Right Arrow	Moves to the right-most region of data
Ctrl + Left Arrow	Moves to the left-most region of data
Ctrl + Up Arrow	Moves to the top region of data
Ctrl + Down Arrow	Moves to the bottom region of data

Faster editing/selections in your Worksheet

ESC	Cancel an entry in the cell or formula bar
Ctrl + Space	Selects the entire column
Ctrl + Space + (+ or –)	Adds/deletes columns (need number pad to use +/- keys)
Shift + Space	Selects the entire line
Shift + Space + (+ or –)	Adds/deletes line (need number pad to use +/- keys)
Ctrl + 1	Display the Format Cells dialog box.
F2	Direct cell edit

Practice – Short Keys:

Open the Excel file**Excel Advanced-English.xls**and spend about 5 minutes trying some of the short keys above in one of the worksheets.

After practising, CLOSE the file **without** SAVING before going on so that so that any changes you may have made to the worksheet don't alter the next activities.

A.2 – Sorting

Sorting is the most frequently used data management function in Excel.

<u>Why Sort your Data?</u>

Since keeping track of your data is a common use of spreadsheets - like records of expenditures, coverage,target population…- it is not unusual to want to change the order of the information (or records) in your worksheet.

Sorting can also allow you to check the extremes or range of your data. For instance, sorting coverage data by province in ascending order will place the province with the lowest coverage at the top of the list, and the province with the maximum coverage at the end of your list.

Finally, the Sort function is extremely useful when merging two different files together that have at least one common column. For example, if you have two files listing all the Districts in your country, but one showing program expenditures by District and the other showing program budget by District. You could sort both files by District name, then copy and paste the expenditure data into a column next to the budget data on one worksheet, allowing you to analyse and compare the data more easily. By sorting the District column in both files you can be sure the correct expenditure data remains in line with the correct District name and program budget data. If you would like an example of this using population and coverage data, open file**Excel Advanced-English.xls**, and go to worksheet **Re-SortCoveragePop.**

<u>Using the Sort Function</u>

When you sort, Excel rearranges rows, columns, or individual cells by using the sort order that you specify. You can sort lists in ascending (1 to 9, A to Z) or descending (9 to 1, Z to A) order, and sort based on the contents of one or more columns.

Practice – Sorting Data:

Open the file **Excel Advanced-English.xls** if you haven't already, and make sure the worksheet **Cold Chain** is showing. You can see that the worksheet shows cold chain equipment for a set of provinces and districts, organised by columns. You also notice that data is first sorted alphabetically by the name of the province, and then by the name of the district.

Short Key Sorting Method
1. It is always prudent to save your file prior to using a function. (If you make a mistake you can always close down without saving and retrieve your original file).
2. Let's say you want to know who has the oldest cold chain equipment so you can begin planning how many refrigerators and freezers may need to be replaced in each of the next 5 years. Click at the top data cell in the column you would like to sort by, for example cell F2. Generally avoid clicking the column title to minimise the chance that Excel will sort it as well.

3. Press **Shift & Space Bar** (at the same time). This will then select the remainder of the row, while still identifying to Excel the column you want to sort by.
4. Then press **Shift End & Down** (at the same time). This highlights the remainder of the cells while still identifying to Excel the column you want to sort by.
5. On the Tool bar click on the Sort Icon (either ascending or descending.)
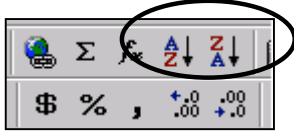
6. All the columns on the worksheet should be automatically resorted according to the year of the equipment, allowing easy estimation of how many of what equipment may need to be replaced in each year.

Warning: Always be cautious when sorting to make sure data remains associated correctly with one another. The technique above should ensure this happens. However, if only some of the columns get sorted, and Excel can be tricky about this sometimes, you run the risk of getting different columns of data out of order.   And remember that you can use the EDIT-UNDO and REDO features to go back and forth after sorting to verify that rows have the correct data.

Regular Sorting Method
Excel allows you to do one or multiple sorting at once. For example, look at the **Cold Chain** worksheet again. You could determine in what year, what source (UNICEF, JICA, etc) was supporting what models of equipment. This might give a program manager an idea of what agency to approach again for new, updated equipment.

To do multiple sorting:

1. Click any cell with data.
2. Go to menu **DATA-SORT**
3. Excel automatically highlights the cells it intends to include in the sort (in this case it should be cells from **A2 to H47.)**
4. In the dialogue box which appears (see below), Select:
    Sort by: Year (Ascending)
    Then by: Sources (Ascending)
    Then by: Model (Ascending)
    The Cold Chain worksheet does have a "**Header**" or title row. Stipulating that were asked avoids including it in the sort.
5. Then click OK
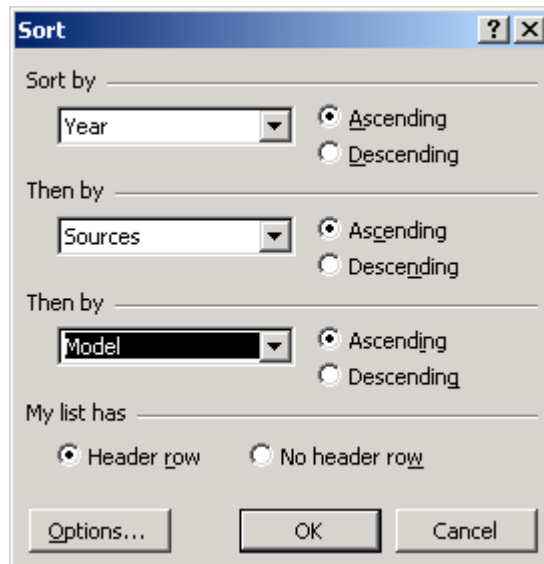
The resorted data should show that for example, JICA bought 3 refrigerators in 1991, and UNICEF bought one, as well as clearly showing the model and quantity.

Warning: Always be cautious when sorting to make sure data remains associated correctly with one another. If only some of the columns get sorted, you run the risk of getting different columns of data out of order.   And remember that you can use the EDIT-UNDO and REDO features to go back and forth after sorting to verify that rows have the correct data.

A.2 – AutoFilter

The AutoFilter is not as well know to most Excel users but is a useful function for quickly viewing your data.

Why use a filter?

Filtering lets you explore data in Excel by allowing you to view the data that meets your chosen criteria. Excel's AutoFilter function makes it easy to hide everything in a database except the records you want to see.

When you choose the AutoFilter command, Excel inserts drop-down list buttons to every cell with a field name in that row. This will allow you to filter all your columns. To add a filter on one particular range of data, first highlight the column of information you wish to filter and then select AutoFilter.

Practice - Creating a Filter:

Open the file **Excel Advanced-English.xls**, and select the worksheet **Cold Chain.**

All you have to do to filter a database is:

1. Position the cell pointer some place on the worksheet (e.g. on **Cold Chain**)
2. Go to menu DATA-FILTER-AUTOFILTER as shown below
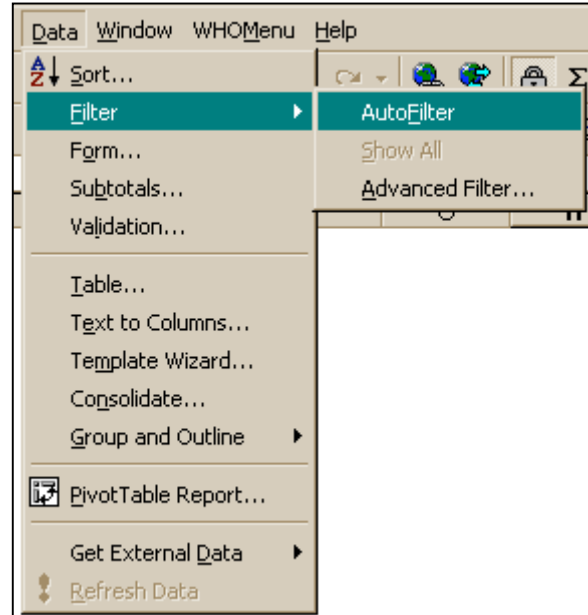
1. Click any columns drop-down list button (top right of each column) to open a list box containing all the entries made in that field and select the one you want to use as a filter. You can note that Excel inserts a drop-down list buttons to every column. This will allow you to do filter one or many columns.

2. Excel then displays only those records that contain the value you selected in that field (all other records are temporarily hidden). The example below shows a multiple filter to view only the cold chain equipment in Makete district that was purchased in 1995. The Filter criteria's were District (selecting Makete) and Year (selecting 1995). Try this on the Cold Chain worksheet.

3. When you want to display all the records in the database again, choose menu DATA-FILTER-SHOW ALL. Alternatively you can also remove a filter from a particular field by selecting its drop-down list button and then selecting the "All" option at the top of the drop-down list.

A.3 – Subtotals

Excel's Subtotal function is an excellent way to summarize data contained in a sorted dataset.

Why use Subtotals?

You can sum, count, average, etc groups of data in your dataset. For instance, you might want to count the number of cold chain equipment you have in each province, and have it subtotal with each change to the data.

Practice: Creating a Subtotal

Open the file **Excel Advanced-English.xls** if you haven't already, and make sure the worksheet **Cold Chain** is showing.

1. Sort the list by the column for which you want to calculate subtotals. In this case, use the DATA-SORT feature to sort the data by Province then District. The dataset must be sorted according to whatever column(s) is used for subtotalling.
2. Click a cell in the list
3. Select menu DATA-SUBTOTAL
4. In the "**At each change in**" box, click the column that contains the groups for which you want subtotal. That way a subtotal will reflect the function below for one province then the next, or for one district then the next. For example, select "**District**"
5. In the "**Use functions**" box, select if you want to sum, average, count. Select "**Count**"….
6. In the "**Add Subtotal to**" box, select the column where you would like the subtotal. Select "**Cold Chain Equip.**"
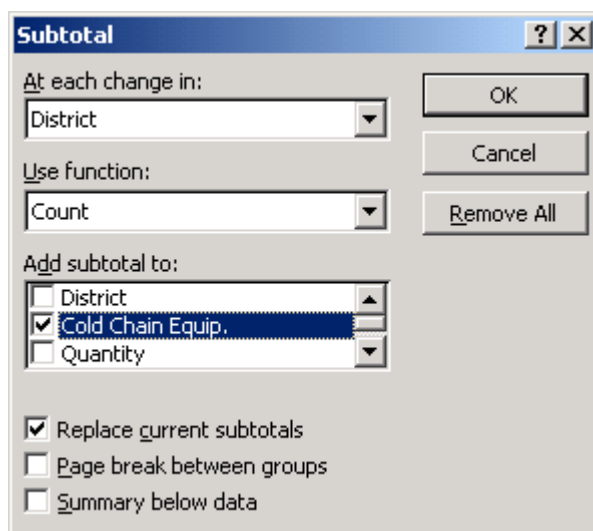
The result of the subtotals are a Grand Total at the top, then subtotals showing the number of items of cold chain equipment listed for each district.

Microsoft Excel - XL_DataManagment.xls

File Edit View Insert Format Tools Data Window Help

A2 =

		A	B	C	D	E	F	G	H
	1	**Province**	**District**	**Cold Chain Equip.**	**Quantity**	**Model**	**Power**	**Year**	**Sources**
	2		**Grand Count**		47				
	3		**Bagamoyo Count**		5				
	4	Arusha	Bagamoyo	Refrigerator	1	RCW42EK	Electric	1992	UNICEF
	5	Arusha	Bagamoyo	Refrigerator	1	MK4010	Electric	1995	UNICEF
	6	Arusha	Bagamoyo	Refrigerator	1	RCW42EK	Electric	1995	UNICEF
	7	Arusha	Bagamoyo	Freezer	1	SB142	Electric	1996	JICA
	8	Arusha	Bagamoyo	Refrigerator	1	RCW42EG	Electric	1996	UNICEF
	9		**Bunda Count**		4				
	10	Arusha	Bunda	Refrigerator	1	RCW42EK	Electric	1991	JICA
	11	Arusha	Bunda	Refrigerator	1	RCW42AC	Electric	1992	UNICEF
	12	Arusha	Bunda	Freezer	1	SB142	Electric	1995	JICA
	13	Arusha	Bunda	Refrigerator	1	MK4010	Electric	1995	UNICEF
	14		**Kilosa Count**		6				
	15	Arusha	Kilosa	Refrigerator	1	MK4010	Electric	1995	UNICEF
	16	Arusha	Kilosa	Freezer	1	SB142	Electric	1995	UNICEF
	17	Arusha	Kilosa	Freezer	1	SB142	Electric	1995	UNICEF
	18	Arusha	Kilosa	Freezer	1	V240KE	Electric	1995	UNICEF
	19	Arusha	Kilosa	Refrigerator	1	RCW42EG	Electric	1996	UNICEF
	20	Arusha	Kilosa	Refrigerator	1	RCW42EK	Electric	1998	UNICEF
	21		**Muleba Count**		3				
	22	Arusha	Muleba	Refrigerator	1	MK4010	Electric	1996	UNICEF
	23	Arusha	Muleba	Refrigerator	1	RCW42EG	Electric	1996	UNICEF
	24	Arusha	Muleba	Refrigerator	1	RCW42EK	Kerosene	1997	UNICEF

A.4 – Pivot Table

Pivot tables are extremely powerful tool to use if your data is normalized. It can save a great deal of time when you are processing your data.

Why Use a Pivot Table?

Pivot tables provide a way to easily summarize, analyze, consolidate, filter, prepare data for charting, and report on large quantities of raw data in a fast, flexible, and ad hoc manner. They are called pivot tables because you can change their layout by rearranging, or pivoting, the row and column headings quickly and easily.

> *Warning: Data must generally be constructed as that we've used (e.g. Cold Chain), with one or more classifications (e.g. Province and District) along one row with additional variables or pieces of data continuing along the row, all pertaining to the original classifications. Such a standardised format for data allows one to take most advantage of Excel's features. This is known as normalized data.*

Practice: Creating a Pivot Table

Creating a Pivot Table is very easy. All you need to do is follow the Pivot Table Wizard.

Open the file **Excel Advanced-English.xls** if you haven't already, and make sure the worksheet **Vaccine Coverage** is showing.

1. Select the first cell of the range (e.g. cell A1) before starting the Pivot Table Wizard, Excel will automatically search down and to the right to find the range's extent and use that range in the pivot table. The easiest option is to click on the upper-leftmost cell of the range.
2. Go to menu DATA-PIVOT CHART REPORT.
3. In the first dialog box, choose "Microsoft Excel list or database", and click "Next".
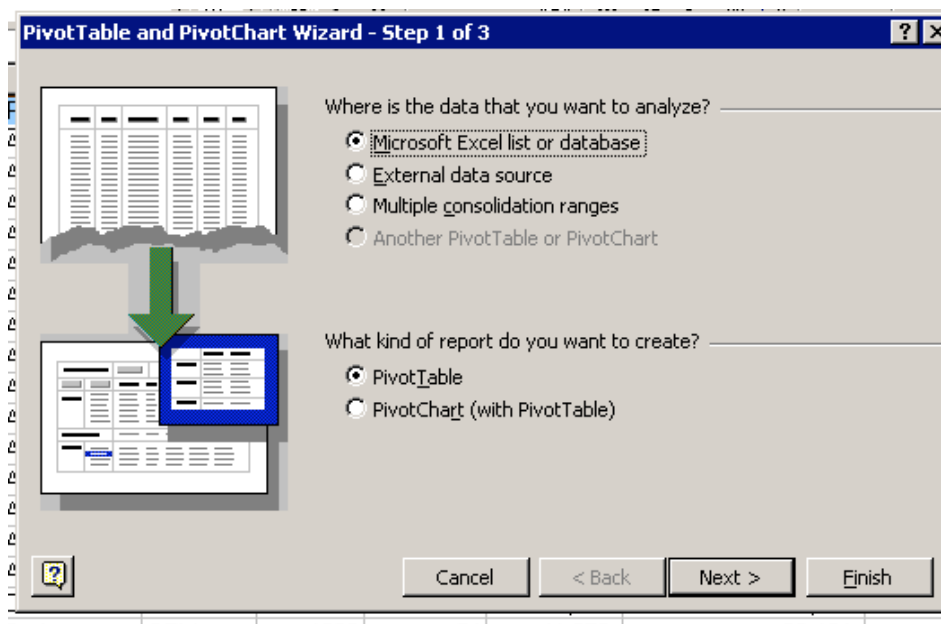
4. In the second dialog box, accept the "Range" by clicking "Next". Excel will help select the range of data you wish to Pivot including headers/titles to you columns.
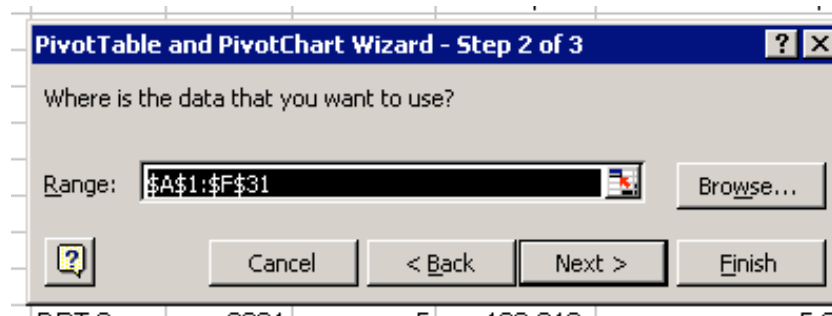
5. The third dialog box should look similar to below.
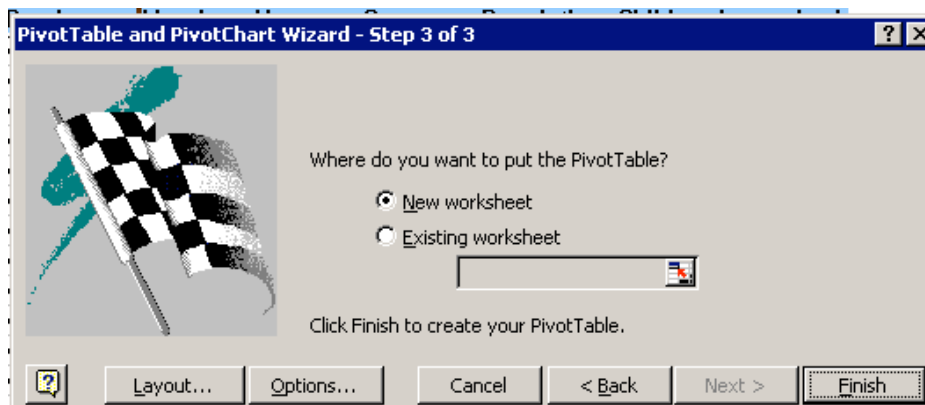
Select where you would like to put your Pivot table and click on Finish. Excel will open a new worksheet which looks like the picture below.
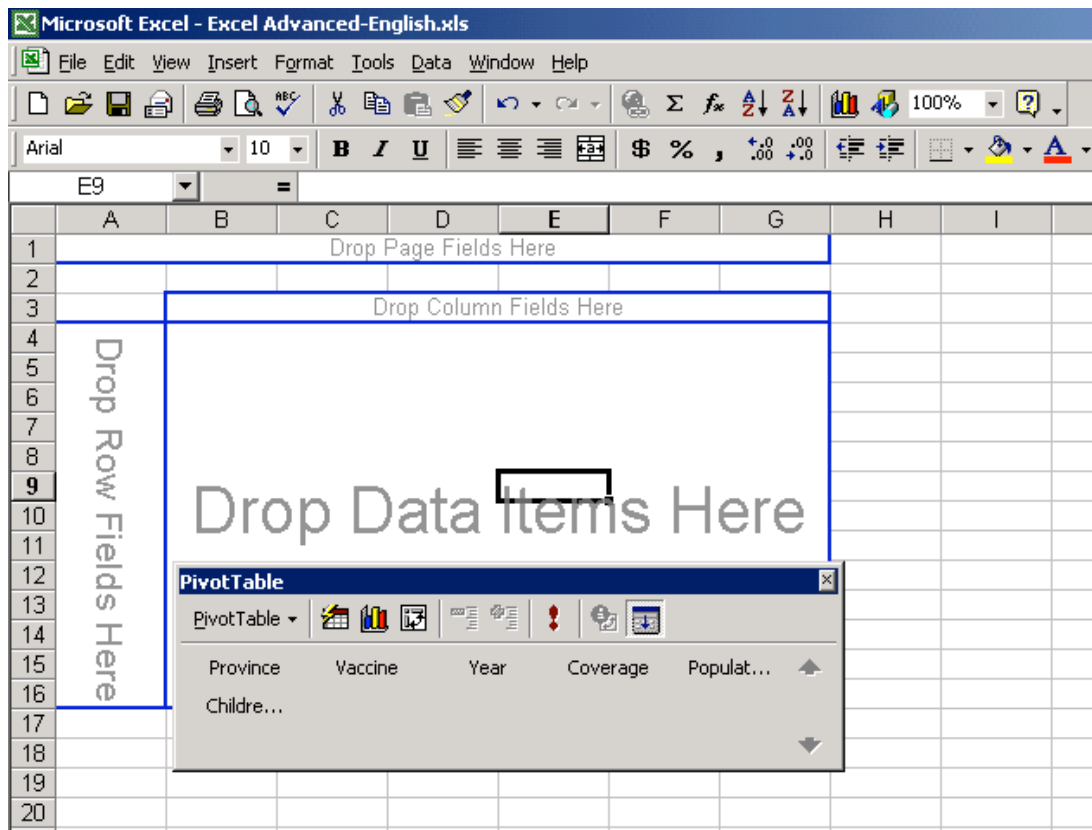
The menu in the middle (PivotTable) shows the titles of all the columns from the original spreadsheet. Behind outlined in blue, one can see areas where the different column headings can be dragged and dropped. Buttons at the right of the dialog box represents the data fields. You drag and drop the buttons to create the desired layout of your table. If you place the wrong field onto the table you can remove it by dragging it anywhere outside the table.

Row: Contains the field(s) to be used as row titles
Column: Contains the field(s) to be used as column titles
Page: Is used to filter the row and column data for a particular Page field value. A Page dropdown will contain all the unique values for the field(s) contained in Page.
Data: At least one field must be placed in the data area. This will be the summary data that is reported inside the rows and columns of the pivot table. The same field can be placed in the data area more than once.

Important: Excel always makes a Pivot Table perform a mathematics function, generally adding or counting what is placed in the data field. Note that numeric fields will default to Sum and non-numeric fields will default to Count for the summary statistic. You can change the default summarization type by double-clicking on the button in the data area and selecting a different calculation type (Sum, Count, Average, Min, Max…).

Let's give it a try. Drag the word Vaccine and drop it on the column bar. Drag the word Province and put it in the row bar. Drag the word coverage and put it in the Data section. You should get the following.

What does it show?

--The top left corner shows that the operation Excel performed was to "Sum of Coverage." Therefore, Excel went to every place where Arusha reported a coverage number for BCG, for example, and added them all up, to 268.7. So it added the 1999, 2000, and 2001 coverage together…the output is totally meaningless. This illustrates that one must be cautious with excel tables, and their interpretation.

Let's look at the data again, but first double click on the top left corner saying "Sum of Coverage." Select "Average of Coverage". You get the table as below.

What does the data show?

It now shows the average coverage for each vaccine, from 1999 to 2001. So the average coverage for BCG in Arusha over these three years was 89%.

The data is reported with 2-4 decimal places. To make it easier to read, select all the cells with numbers, then push the "," button on the toolbar, then press the button to decrease decimal places.

Notice that a filter is available on the year (shown as "All"), so aggregate data for one or more years can be shown. Notice also that by default Excel will provide you with Grand Totals, although in some cases they may not be relevant or meaningful. The grand totals could be particularly useful when using Pivot Tables for financing data.

You can add additional fields into the row or column headings. Give it a try, but be cautious as the data can get pretty complex to interpret. There's a sample shown below.

If "Vaccine" had been put where "Year" was placed, on the "Page" section, it would have allowed a similar sheet, but showing one or all vaccines for 3 years. If you have time, try to construct such a Pivot Table.

While vaccine coverage data was used to illustrate pivot tables, they cab be quite useful for tracking, presenting, and analysing budget and expenditure data over time.

B – DATA ANALYSIS

The analysis ToolPak that is available in Excel is a useful "Add-in" to activate. As an "Add-in" you must specifically activate for it to be available. It will allow you to do descriptive and more advanced statistics at the click of the mouse. For example it quickly gives you the mean (average), median, statistical difference, and other basic statistics for a set of numbers. These types of analysis don't pertain as frequently, possibly, to financial analysis, but are likely useful as part of general program management and for interpreting immunisation data.

Enabling Analysis ToolPak
If the TOOLS menu does not have an option for DATA ANALYSIS, you need to install the Analysis ToolPak in Microsoft Excel.

To install the Analysis ToolPak

1. On the TOOLS menu, click ADD-INS.

2. Select the Analysis ToolPak check box.
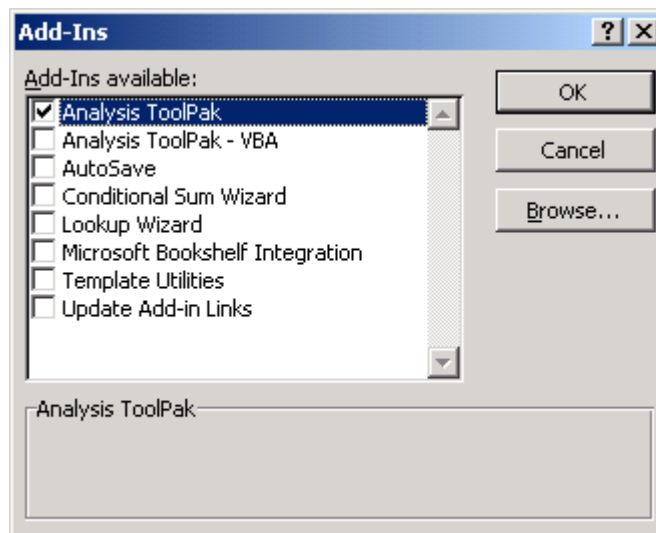
Practice: Analysis ToolPak
Open the file **Excel Advanced-English.xls** if you haven't already, and make sure the worksheet **Vaccine Coverage** is showing.

Example: Analysing the recent Measles coverage for Arusha Province.

1. Before using an analysis tool, you must arrange the data you want to analyse in columns or rows on your worksheet. This is your input range. Use the SORT command to sort the table by Province, then Vaccine, then Coverage. Remember you can do this all in one step using the SORT command.
2. On the TOOLS menu, click DATA ANALYSIS.

3. In the Analysis Tools box, click "Descriptive Statistics."
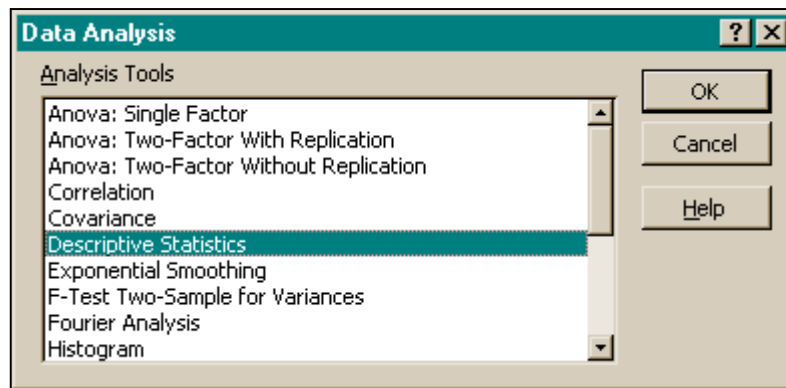
4. Enter the input range and the output range by highlighting the numbers you want analysed, and then select the options you want (See below).

5. When you click "OK," the results are presented in a table, similar to below:

Column1	
Mean	43.00
Standard Error	15.95
Median	52.00
Mode	#N/A
Standard Deviation	27.62
Sample Variance	763.00
Kurtosis	#DIV/0!
Skewness	-1.31
Range	53.00
Minimum	12.00
Maximum	65.00
Sum	129.00
Count	3.00

This shows, among other things, that over the past three years in Arusha Province, Measles coverage has averaged (mean) 43%, the median was 52%, and the minimum and maximum reported coverage was 12% and 65% respectively.

Conclusion
In this module, we worked through some advanced Excel skills. Mastery of a program like Excel takes much time and use. However the skills covered in this are some of the most commonly used and useful for financial and programmatic analysis.

And when things don't seem to be working correctly, don't forget to try the excellent HELP menu. Just type your question and Excel will help try to find the item you are interested in.

Some other skills you may want to investigate if you are not familiar with them:

Dragging Formulas: Excel is excellent for inserting a number or formula into one cell, then "dragging" it in any direction to apply to other cells. This is a great way to save time. Note that using a "$" in a function means that the number will not change when moved, or drug to another cell.

Logical Functions (See below)

True/False: These can be verify powerful ways of validating data. For example, a cell can be told to return a "FALSE" statement if certain conditions aren't met in other data. This can be used to verify validity of data.

IF: The IF function can be used to apply a set of conditions for a total to be returned, and what should be given, if those conditions aren't met.
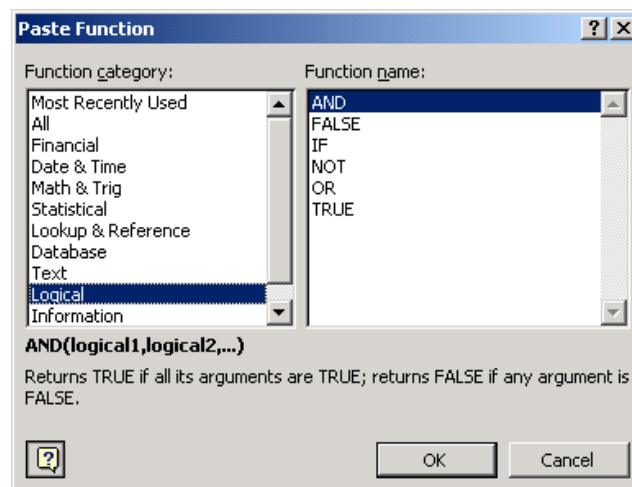
Paste Link: Sometimes the same piece of data, cost or number should be reflected in more than one place on a worksheet, or even on more than one worksheet. This is an ideal time to use PASTE SPECIAL - PASTE LINK. Then, when one changes the original cell, the change is automatically reflected in all the other linked cells. Clicking on the linked cell shows that it has only the location of the source cell (e.g. "=B5".) This can be particularly useful if for example you have a currency exchange rate you are applying to multiple numbers. If all the calculations link back to one cell, adjusting that one cell with a new exchange rate automatically adjusts all the linked data.