Instruction for Data Analysis with EXCEL

Overview of Excel Spreadsheets

We begin by reviewing some basic concepts for spreadsheet programs, starting from the full program and gradually zooming in to the details. But first, open the Excel *Workbook* file named "*Excel Data Analysis"* and save your *Workbook* with a new name by selecting **File>Save As** from the Menu Bar and giving the file the name "*Your_Last_Name-Excel Data Analysis*".

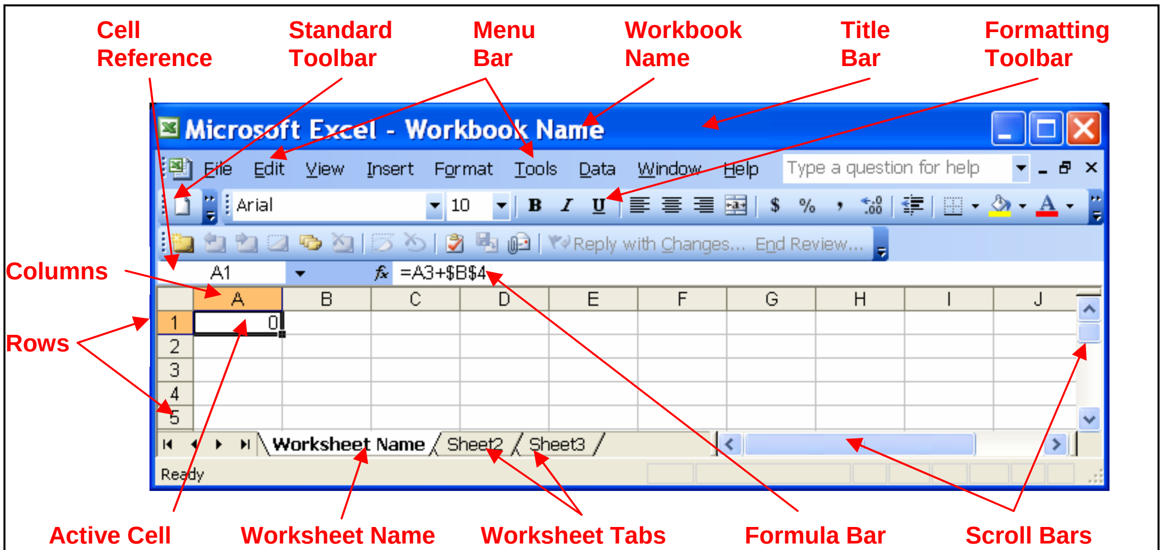

Figure 1 shows a screen shot of an Excel *Workbook* file and identifies the key components.

Spreadsheet: A spreadsheet program is a sophisticated and versatile analysis tool and a display tool for numeric, text, and graphical data. A spreadsheet consists of a two dimensional array of boxes (referred to as cells) into which text, symbols, numbers or formulas can be typed. Data in the spreadsheet cells can be dynamically manipulated, linked together and presented in tabular or graphical form. This means that changing the value of the data in one or more cells can affect the values displayed in other cells in the spreadsheet.

Workbook: A collection of *Worksheets* saved together as a single spreadsheet file.

Worksheet: A two dimensional collection of cells. A *Workbook* can have multiple *Worksheets*, each identified by a name on the tab at the bottom of the spreadsheet window. See Note 1 for an example.

Cell: A single box containing data. The cell is identified by its column and row, for example as A1 in Figure 1. Cells can be identified in several ways:

> *Relative Reference:* A cell name that will change when the cell formula is moved to another cell. The format is A1. See Notes 1 and 4 for examples.
> *Absolute Reference:* A cell name that will remain fixed when the cell formula is moved to another cell. The format is $A$1. The proceeding $ "fixes" the column or row. See Notes 1 and 4 for examples.
> *Cell Range:* A block of cell can be denoted with the format A1:C5, where the two cell references separated by a colon are the upper left and lower right corner cells of a rectangular block of cells. See almost any Note for an example.
> *Worksheet Reference:* A cell from another *Worksheet* can be identified by extending its reference to include the *Worksheet* name, for example as Worksheet_Name!A1 . The trailing ! signals that Worksheet_Name is a *Worksheet* name. See Note 1 for an example.

Named Reference: A cell can be linked to a name or variable that can be used in formulas. See Note 3 for an example.

Data: Data can be input into a cell in several ways:

Keyboard: Highlight a given cell and type either text or numeric data. See Note 1 for an example.

Cut and Paste: Highlight a cell and either **Cut** (Cntrl-X) or **Copy** (Cntrl-C) the contents. Highlight another cell and Paste (Cntrl-V) the data. Data can also be Cut and Pasted from other programs besides *Excel* using the *Windows* clipboard. See Notes 8 and 10 for examples.

Highlight and Drag: See Notes 2 and 4 for examples.

Transfer by Formula: Data can be automatically inserted in a cell by using a formula (see *Formulas* below). See Note 1 for an example.

Importing Data: Data for a single cell or for a block of cells can be inserted from an external file using the **Import Data** command. See Note 2 for an example.

Formulas: A formula calculates the contents of a given cell by performing mathematical, statistical or other operations on the contents of other cells. The formula for a given cell is displayed, and can be entered or edited, in the window of the Formula Bar. A formula always begins with an equals sign, "=".

Cloning Formulas: The formula from one or more cells can readily be duplicated into other cells. The way the formulas are copied and modified in the new cells depends of the type of cell references used in the formula in the original cell(s). See Notes 4 and 5 for examples.

Built In Functions: Excel has a wide array of built-in functions that can be used in formulae. A complete list of available functions, along with function descriptions and syntax and a search feature, can be accessed from the Menu Bar command Insert>Functions. See Note 5 for an example.

Formula Auditing: A clever set of tools to troubleshoot formulae for a given cell are accessed from the Menu Bar command **Tools>Formula Auditing**. See Note 4 for an example.

Moving Cells:

Cut and Paste: Highlight a cell and either **Cut** (Cntrl-X) or **Copy** (Cntrl-C) the contents. Highlight another cell and **Paste** (Cntrl-V) the data. See Notes 7 and 10 for examples.

Highlight and Drag: Highlight the cells you want to moves, moving the cursor to the edge of the highlighted box until it switches to the shape of a vertical and horizontal arrow, and then dragging the highlighted box into place. See Note 2 for an example.

Insert and Delete Rows and Columns: To add or remove whole rows or columns of cells, highlight a row, column or cell, select **Insert>Rows** (or right click and choose **Insert** or **Delete**) and choose the position. See Note 2 for an example.

Formatting Cells or Numbers: Select a cell(s), right click and select **Format Cells**. There are many options for the number format, text alignment, font, border, cell background, and other things.

Plots: To add a plot, click on an empty cell and select **Insert>Chart** to bring up the **Chart Wizard**. More details on creating and editing chart are given in Notes 7 and 8 below. To position a chart, highlight it and drag it into place. To change its size, highlight the chart and drag the corners to the desired size. To edit the chart, highlight the chart and right click on either the chart, a data point or curve, an axis, a label, or other graph components.

Instructions for Completing the Excel Data Analysis Spreadsheet

What follows is a detailed, step-by-step, set of instructions to complete the *Excel Data Analysis Spreadsheet*. The intension is to teach many of the basic capabilities of *Execl* for scientific computing by example. Open the spreadsheet, if it is not already open. Begin in the *Data & Analysis Worksheet*. Follow the instructions below for each sequentially numbered Note #, beginning at the location of the Note # label in the spreadsheet. [*Hint:* You need not type the quotes "" in the instructions below.]

Note 1
Fill In Header: Highlight cell L1 by clicking in the cell. Type your name in cell L1. Hit **Enter**, **Tab** or click outside the cell to complete the entry. Repeat this for the current date in cell K1. These are examples of direct keyboard data entry. Go to the *Experiment Description Worksheet* by clicking on its tab at the bottom of the page. Note your name and the date now appear in cells C1:C2 without you having entered them. Click on cell C1 and look at the entry in the **Formula Bar**. It sets this cell's content equal to the content of cell L1 in the *Data & Analysis Worksheet*. These are examples of an absolute reference, a worksheet reference and a cell range, as well as an example of transfer of data by a formula.

Note 2
Data Entry: Insert the time and displacement data from your data file. Position the cursor in cell A14 and select **Data>Import External Data>Import Data** . Highlight the data file using the dialog box. In Step 1 of 3 select the **Start Import Row** to omit any headers in the data file. In Step 2 of 3, select **Tab delimited data**. Select the defaults to complete the rest of the dialog box. The time and displacement data should now be inserted in the proper cells. [*Caution:* Sometimes the data are inserted in two new columns to the left of the original column **A**. Move the data into the cells under the "**Time (s)**" heading by highlighting all the data, moving the cursor to the edge of the highlighted box until it switches to the shape of a vertical and horizontal arrow, and then dragging the highlighted box into place. Finish by highlight the two now blank columns at the left by clicking on the headers "**A**" and "**B**", right clicking, and selecting delete.]

Note 3:
Define The Fitting Function: The fitting function is listed in cell J14. You must now define the rest of the fitting parameters, as named in cells K18:K21. Enter the symbols w, θ and yo (as used in the foluma in cell J14) in cells L19:L21. Highlight cells K18:L21 and select the **Menu Bar** command **Insert>Name>Create** . Enter values 1, 2, 3 and 4 in cells L18:L21, respectively. Now the name "A" from cell K18 has the numeric value 1 from cell K18 whenever "A" is used in a formula; the same is true for the other three parameters named in cells J19:J21. To finish this part, add units "m", "Hz" and "m" to cells M18, M19 and M21, respectively. (Phase θ defined in cells K20:K21 has no units.)

Note 4:
Calculate Other Columns: Insert a value of 0.002 in cell C14 for the first uncertainty in the displacement data. Enter the formula "=A*sin(w*A14+ θ)+yo" in cell D14 for the first fit value. Enter the formula "=$L$21" in cell E14 for the first offset value. Enter the formula "=D14-B14" for the first deviation in cell F14. Enter the formula "=(F14/C14)^2" for the first weighted deviation squared in cell G14. And now for the spreadsheet magic! Highlight cells C14:G14. Grab the small black square at the lower right of the rectangle, hold down the right mouse button, and pull it down so that all the cells down to the bottom of the data entry in columns A and B are highlighted. When you release the mouse button, the values of all the cells are instantly calculated. Click on some of these new cells. When *Excel* cloned the formulas, it automatically incremented the cell values so the each cell was calculated with the correct data values. The values in the cells in columns C and E for the Error and Offset are constant because only numeric values and absolute cell references (ones with $ in the cell names) were used in the formulas. The values in columns D, F and G are different in each cell, based on the relative references for a given row.

Even better, change the value in cell L18 assigned to the variable A and watch all the values that are calculated using A in the Fit, Deviation and Wgt. Dev. Sq. columns change in response to the new value for A. Pretty nifty keen, huh! This is what makes spreadsheets like *Excel* so handy for "what if analysis", not unlike *Mathcad*. However, spreadsheets are not as transparent as *Mathcad*, since the formula are hidden and often hard to troubleshoot (see below) and do not have the flexibility for calculations and documentation that *Mathcad* has. Still, the simplicity of spreadsheets certainly can have its advantages.

Just for fun (a geekish sort of fun, but fun none the less), try playing with the cell auditing procedure, a clever set of tools to troubleshoot formulae. Highlight cell D14 and select the command **Tools>Formula Auditing>Show Auditing Toolbar** from the **Menu Bar**. Try the buttons to **Trace** and **Remove Precedents** and **Trace** and **Remove Descendants**. This clearly has a high geek factor, and can even be useful.

Note 5:
Calculate Data Statistics: We will use built-in functions to complete this section. In cell A5 enter the formula "=AVERAGE()". Position the cursor between the parentheses in the Formula Bar, then highlight all of the time data in column A below. A range should appear between the parentheses, such as "=AVERAGE(A14:A55)". Highlight cell A5 and clone it into cells A6:A10. Now edit the formula for cells A6, A7, A8 and A10, replacing the function AVERAGE with STDEV, MIN, MAX and SUM, respectively. For cell A9 enter the formula "=A8-A7". Now highlight cells A5:A10 and clone them into the region B5:G10.

Note 6:
Calculate Goodness of Fit: Enter the number of time and displacement data points in cell K5 (you have to count them, since each data set varies in length). Enter 4 in cell K6 for the number of fitting parameters (A, w, θ and yo). Enter the formula "=K5-K6" in cell K7 for the degrees of freedom, which equals the number of data points minus the number of fitting parameters. Chi squared is just the sum of the weighted deviations squared (which was already calculated in the data statistics for column G), so enter the formula "=G10" in cell K8. The reduced chi squared value is just chi squared divided by the number of degrees of freedom, so enter a formula "=(G10/$K$7)" in cell K9.

Note 7:
Main Graph: To generate a graph of displacement versus elapsed time, highlight cell I23 and select the **Menu Bar** command **Insert>Chart** . Choose an **XY (Scatter)** chart type with data points connected by lines. Highlight the **Series** tab in the next dialog box and select **Add**. To identify the **X values** to plot, click on the icon at the right of the **X Values** window, highlight the time data value cells, and press **Return**. Repeat this for the **Y Values**, selecting the Displacement data cells. Name this "Data" instead of the generic Series1. To add a second line to the graph, select **Add** again. Name Series2 "Fit", select the same time values for the **X Values** and the Fit cells for the **Y Values**. Add a third series named "Offset", with the same **X Values** and the Offset cell as **Y Values**. Select **Next>** to go to **Step 3 of 4-Chart Options**. Under the **Titles** tab add a graph title, an x axis label, and a y axis label. Under **Gridlines** tab, deselect all gridlines. Under the **Legend** tab select bottom. Select **Next>** and the **As object in:** button. Select **Finish**. Your graph should appear in the *Worksheet*. The final formatting can be done by clicking on the individual items to format. Click on a data point to display the Patterns tab of the **Format Data Series** dialog box; select **None** for **Line** and you favorite symbol and size for **Marker**. Do the same for the Fit series, selecting no markers and a solid line. Do the same for the Offset series, selecting no markers and a dashed line. Click on the vertical axis. Select **Inside** for the major and minor tick marks under the **Patterns** tab and adjust the plot range as needed under the **Scale** tab. Under the **Font** tab, deselect **Auto scale** and choose a 10 point regular font. Do the same for the horizontal axis. Click on the **Value (X) Axis Title**, deselect **Auto scale**, select a 10 point regular font, and drag the label to the place you want. Repeat this for the **Value (Y) Axis Title** and the graph **Title** (use a 12 point font.). Make any final adjustments to make the graph appearance the same as the example.

Note 8:
Residual Graph: Highlight the Main Graph, select copy, and paste a copy in a cell just below the Main Graph. Highlight the new graph and drag to reduce the graph vertical height. Select and change the **Title**. Select **Source Window** and change the **Y Values** of Series1 and Series 2 to the Deviation and Error cells, respectively. Remove Series3. Make any final adjustments to make the graph appearance the same as the example.

Note 9:
Experiment Description Worksheet: Select the *Experiment Description Worksheet* tab at the bottom of the window. Add text in cell A4:A8 to describe the experiment. Highlight cells A11:A15 and select the **Menu Bar** command **Format>Format Cells>Alignment**. Select the **General Horizontal** alignment, the **Top** vertical alignment, and **Merge cells**. Select **OK** and enter text describing the equipment in the new box. Add a photo of the apparatus in cell A18 if desired. Make any final adjustments to make the *Worksheet* appearance the same as the example.

Note 10:

Project Report: Select the *Project Report Worksheet* tab at the bottom of the window. Copy cells I1:J2 for the name and date from the *Experiment Description* window to cells I1:J2. Copy cells A3:A8 for the experiment description from the *Experiment Description* window to cells A3:A8. Merge and format cells A12:A18 to accommodate text summarizing the results of your experiment. Merge and format cells A21:A25 to accommodate text summarizing the fitting. From the *Data & Analysis Worksheet*, copy the Fitting Function cells J13:J14, Fitting Parameter cells J16:M21, the Goodness of Fit cells J4:K9, and the two graphs; paste and position these in the *Project Report Worksheet.* Make any final adjustments to make the *Worksheet* appearance the same as the example.

Review, save and submit your final *Excel Workbook* by selecting **File>Save** or **File>Save As** from the **Menu Bar** and giving the file the name "*Your_Last_Name-Excel Data Analysis*".