

Using Microsoft Excel

In many of your courses, especially science courses, you will be working with numerical data. At times, you will be working with a LOT of numerical data. Although you are welcome to churn through this data with paper and pencil (and maybe a calculator), computers are made to crunch numbers so it will probably be easier to use these valuable tools when your data sets get large.

Entering data:

Open a new Excel workbook. Enter the following into separate cells in the workbook without the “{}”: {chem}, {150}, {3rd}, {7.216}, {18.31g}, {‘3.18}

How did the program format (left, right, centered) what you entered in each cell? What does this difference tell you about data you enter?

Formatting data/cells:

Enter the following into separate cells without the “{}”: {2.816}, {1.2000}, {6543210123456}, {6500000000000}, {0.00081}, {0.0000000018}

Do the cells display the numbers as you have entered them? How have they changed? Is this change important?

Point to one of the cells and right-click. From the pop-up menu, select the “Format Cells...” option. *How is the cell currently formatted? Is there a better option for formatting some of the cells to display the information as you entered it?* Format all of the cells to correctly and accurately display the data you have entered.

Enter the following into separate cells without the “{}”: {data}, {point}, {experiment}, {Time of day}, {Checking account balance}

Are these cells displaying the information clearly? What happens to text that is too long to fit in the cell? Click-and-drag the border of one of the column headings to re-size the column width. Double-click the right border of another column heading. *What happened? Are your column widths reasonable?* Right-click in the cell of one of the longer cells and select “Format Cells...” again. Click on the “Alignment” tab and select “Wrap text”. Click “OK”. *Did the information in the cell change?* Click-and-drag the border of the column heading. *What happens to text that is too long to fit in the cell?* Format all of the text cells appropriately.

Using formulas to enter numerical data:

Click on the “Sheet2” tab at the bottom of the MExcel window to bring up a new sheet in your workbook. Let’s say you are going to collect a series of data and you want to number that data. In cell A1, type {Experiment Number:}. *How did Excel format this cell?* Format appropriately. In all other columns, use the first row for a label and cells 2-11 for data.

In cell A2, type {1}. Depending upon the size of your data set, you could proceed to type {2}, {3}, etc down column A, but this quickly becomes inconvenient for large data sets. Instead, we will use a formula. All formulas start with an equal sign. In cell A3, type {=A2+1} and hit enter. *What is the result? Is it correct?* If so, select cell A3 and “copy” it. This can either be done with the mouse and menus, or by holding down “control” and typing “c”. {Ctrl-c}. Select cell A4 and paste either with mouse and menus or {Ctrl-v}. *What is the result?* Use the mouse to highlight cells A5 through A21 and paste again.

Generating Data:

Get a meter stick and a small rubber ball. Bounce the ball a few times to get an idea of its behavior. You are making **observations** about the ball at this point. We would like to do an experiment that quantitatively explores the “bounciness” of the ball you are using. *When the ball is dropped from a given height, how high do you expect it to bounce? Will the height (or relative height) of the bounce depend upon the height from which the ball was dropped?* These are the questions you should address in your experiment and should be the basis of the **hypothesis** for your experiment. **Insert a text box in your data collection spreadsheet and write a formal, testable hypothesis for your expectation in this experiment before you continue.** A formal hypothesis should not simply be a question; it should give some idea of the experiment that will be performed and it should predict the outcome. Since this is a *quantitative* experiment, the hypothesis should also be quantitative. The hypothesis “When the rubber ball used in this experiment is dropped from a height of 95cm, it will bounce” is not a useful hypothesis for this experiment because it is *qualitative* and rather trivial; the ball will bounce, but there is no

indication or estimate of how high it will bounce. Similarly, the hypothesis “When the rubber ball used in this experiment is dropped from a height of 1.00m, it will bounce 2.5m” isn’t very useful because it is obviously false; you know from your initial observations that the ball will not bounce higher than the height from which it is dropped. After you have written a reasonable hypothesis, perform a series of experiments to test your hypothesis and record your observations in the spreadsheet. The data you collect might vary a little bit depending upon your hypothesis, but for any experiment you perform it will probably make sense to label column B “Drop height” and column C “Bounce height” to record those values for each experiment. You should perform 20 individual experiments. Record any other data about the ball you are using that you feel might be important.

In column D, we would like to calculate the difference between the drop height and the bounce height using a formula. Label the column appropriately in cell D1. In cell D2, type $\{=B2-C2\}$ and hit enter. Copy this formula down for each row of data. In columns E and F, we would like to calculate the *relative* bounce height and the *relative* difference between drop height and bounce height as percentages in both cases. Label the columns in cells E1 and F1 and insert formulas to calculate the percentages. *Are your results consistent with the hypothesis you made before performing the experiment?* Check with other groups in the lab who used the same type of ball and different types of balls. *Are their results consistent with yours?*

Calculating statistical data:

One of the biggest advantages of using a spreadsheet is that it is quite easy to calculate statistical information for large data sets. We will calculate statistical functions for columns B-F below the data block. Use column A to label each row or statistic.

$=count(range)$ → To count the number of values, in cell B23 enter the function $\{=count(B2:B21)\}$.

$=sum(range)$ → This adds up the values in the specified range, not always a useful number but we will include it.

In cell B24, enter the formula $\{=sum(B2:B21)\}$.

$=average(range)$ → We could calculate the average with a simple addition and division formula or by using the $\{sum\}$ and $\{count\}$ formulas, but Excel has already combined these formulas for us. In cell B25, enter the formula $\{=average(B2:B21)\}$.

$=max(range)$ → Excel can find the maximum value in a range of cells. In cell B26, enter the formula $\{=max(B2:B21)\}$.

$=min(range)$ → Excel can also find the minimum value in a range. In cell B27, enter $\{=min(B2:B21)\}$.

$=stdev(range)$ → Standard deviation indicates the distribution of a data set. Standard deviation tells us very little about a data set unless it is compared to the average or accepted value of the data being measured. To calculate standard deviation, enter the formula $\{=stdev(B2:B21)\}$ in cell B28. For sufficiently large data sets (more than 4 or 5 data points), standard deviation can be used as the error in a measurement. In cell B29, calculate the percent error with the formula $\{=(B28/B25)*100\}$.

Copy all of these statistical data to columns C-F. Use these results to evaluate your hypothesis. Do your experimental results support or refute your hypothesis?

Sorting data:

From the “View” tab, “Window” box”, open the “Freeze Panes” pop-down menu. Select “Freeze top row”. *What happened to the sheet?* Scroll down. *What happens when you scroll down the sheet?* This can be very helpful if you have very long data sets. Right now, your data is sorted by the order in which you did your experiments. For some presentations, it might be preferable to sort by the actual data. Click-and-drag to select cells A1:F21. (Do *not* select the statistical data at the bottom of your data block.) From the “Home” or “Data” tab, select “Sort”. Use a custom sort to rearrange your data. Sort by one of the height columns (B or C). *What happened?* Undo the sort. Select the data in the “Experiment Number” column and copy the data. In the “Clipboard” pane, pop down the “Paste” menu and select “Paste values”. *What happened?* *Are your sequential point numbers still formulas?* Check them. Re-do the sort.

Creating and using graphs:

When working with numerical data, it is often useful to present the data in a graph. A properly formatted graph can convey a wealth of information in a very compact package and lead the reader through an elegant interpretation of what would otherwise be complicated tables of data. Similarly, a poorly formatted graph can make it *more* difficult for the reader to understand the data being presented, and can lead to incorrect interpretation of that data. “Good” graphs share a number of features:

1. Correct selection of graph type. Excel tries very hard to “think” for you, but like all computer programs, Excel does not always (or even often) select the correct defaults. You must *always* make sure that the graph you intend to make is the one that Excel *actually* produces. The vast majority of the graphs you make in science classes will be X-Y scatter plots. You will almost *never* use what Excel calls a “line graph”.
2. Correct selection of “x” and “y” data. The “x” variable is the independent variable, which means it is usually the one over which the experimenter has some control. The “y” variable is the dependent variable and its value is simply measured. Say you are measuring the volume of a gas at a number of temperatures. As the experimenter, you will control the temperature of the experiment and measure whatever volume the gas happens to occupy; temperature would be the “x” variable and volume would be the “y” variable.
3. The data points in the graph should fill the area of the graph. The “x” and “y” axes should be adjusted appropriately. If all of the data points have “x” values between 23 and 28, it doesn’t make much sense to have a graph in which the “x” axis runs from zero to 50. For some presentations, it is important to show the origin {the (0,0) point} on the graph, but not always.
4. The axes should be clearly labeled, including units, and the graph should have a title.
5. If multiple data series are displayed on a single graph, there should be a clear legend.
6. The graph should be made as large as practical in the space provided. If a graph is going to be on its own separate page, then that graph should *fill* the page. It shouldn’t be a little postage stamp in the corner of a blank page.
7. Trends in the data should be indicated with fit lines, but the individual data points should (usually) *not* be connected. A trend line should be chosen for physically and mathematically legitimate reasons, not just because it “looks right”.

Click-and-drag to select the drop height data (column B). Hold down the “Ctrl” key and click-and-drag to select the bounce height data (column C). Using the “Ctrl” key allows you to select non-adjacent sets of data. From the “Insert” tab, insert a scatter chart. By default, the chart will be inserted in your sheet. To make it easier to format and work with the chart, right-click in the chart window and select “Move chart”. Select “New Sheet” and click OK. In the Chart Tools “Layout” tab, add axis labels and explore other formatting options to make the graph look “good”. *Is there a trend in this data? Can you offer an explanation of this trend?* From the “Chart Tools/Layout” menu, add a linear trendline to this data. Having a trendline is not often very useful without an equation to describe that trendline. Point to the trendline and right-click to “Format trendline”. Choose the boxes to display the equation and the R^2 value on the chart. *What does the equation of this line tell you? Does it make sense? What does the R^2 tell you? Does this linear trendline accurately represent the data you have collected?*

Note on graphs: Although we will often be graphing linear data, you will *almost never* want to use what MSExcel calls a “Line” graph. When a graph is incorrect, it is almost always because it was formatted as a “Line” graph instead of an “X-Y Scatter” graph. Do not use line graphs. For our experiments, a “Line” graph is never correct. Never, never, never. Line graphs are very nearly useless in any class. Do not use them. Is that explicit enough?

What to submit for this lab?

Save your MSExcel file with the data table and properly formatted graph to either a flashdrive or your D2L Locker. You can also email the file to yourself. In any case, upload the file to the “Using Microsoft Excel” dropbox in D2L before you leave lab. *For the purposes of this special lab exercise only* it is acceptable (and even expected...) that your and your partner’s submissions are identical; however, be sure that you both submit the assignment separately in D2L. You may print your data sheets and graph if you like for your own records but the physical copies should not be submitted to your instructor. Only the D2L submissions will be graded. See the grading rubric for this exercise on the next page.

Printing worksheets:

Here are a few hints for how to make your printed worksheets (if you choose to print them) easier to read. Many of these options are found in the “Page Layout” tab. First, print the gridlines that appear on the screen. Large blocks of numerical data can turn into a sea of numbers without gridlines. It is also often convenient to make the data fit a single page width and/or length. If the data will be more than a single page long, you may want to repeat the column headings at the top of each page. Microsoft default margins are almost always much larger than necessary, make them a bit smaller to fit more info on the page. This document is formatted to print 2-sided using single-spaced 11-pt Times New Roman with 0.5” outer margins and 1” inner margin to allow space for a 3-hole punch. Using MSWord defaults this document would be quite a bit longer. Think before you print.

Summary:

This exercise is meant to introduce a number of the features found in MSEExcel, but it is by no means a comprehensive introduction or tutorial. Different disciplines and different instructors will have different needs and different preferences for how to present data and graphs. Always be sure to check with your instructor to determine what is considered “correct” formatting of data and graphs. The only way to become proficient with any program is to use it, explore new features, make mistakes and correct them.

Grading Rubric:

Generally rubrics are provided for your benefit as you prepare your hand-ins outside of lab. Since the assignment for this exercise is being submitted at the end of the laboratory period the rubric is provided so that you are not surprised when your assignment is graded. As you can see, all of the things for which points are being earned are things that were clearly indicated in the directions.

	0 pts	1 pt	2 pts	3 pts	4 pts	
Hypothesis	No hypothesis	Hypothesis is there but not well stated.	Reasonably stated hypothesis			2 pts
Data Sheet		½ point deduction for each poorly formatted item.	Meaningful column labels and widths. Table is easy to follow.			2 pts
Graph	If the graph is not an x-y scatter plot, the graph earns zero points.	Graph is an x-y scatter plot but does not meet any of the criteria noted in the last column.			Graph fills the page and is scaled properly. 1 pt. Correct axis labels and graph title. 1pt Best fit line and equation with R^2 value. 1 pt.	4 pts
Conclusion	No conclusion	Poorly supported conclusion	Conclusion is supported by data presented in table and graph.			2 pts
Total						10 pts