**Spreadsheets in Physics Lab**

**(One week, At-Home)**

**Notes:**

1. This is a one-week lab.
2. This assignment can be done at home, from a campus computer, or from any other computer with Microsoft Excel and Word (or equivalent programs) installed.
3. We recommend doing the lab early in the week, rather than waiting until it is almost due. If you have computer trouble, you will want to have plenty of time to fix it before the deadline. No excuses!
4. Most labs this semester can be done with a partner, but please do this exercise on your own.
5. Getting help:
	1. Your lab TA can answer questions by email or during their office hours listed in the syllabus.
	2. You can also ask advice from lab partner(s) and/or other students.

**Objectives of this lab:**

This lab will introduce you to using *Microsoft Excel.* Many of you will know some of the skills already, but we will definitely use more of Excel than most people know. We will use it during the semester to analyze data, make graphs, do calculations, and fit functions to data.

Note: Excel is widely used in many jobs, so the skills you learn here (and practice all semester) are definitely relevant to almost all careers.

**Equipment:**

A computer equipped with *Microsoft Excel* and *Microsoft Word*, or equivalent software.

**For a free download of Excel (and the rest of MS Office) for your computer (Mac and Windows), go to** [**iuware.iu.edu**](http://iuware.iu.edu/) **and search for *Office 365.*** This will give you a free student license for installing the software on your personal machine. Note that as an IUPUI student, you also have free access to the cloud-based version ([see this page from UITS to get started](https://uits.iu.edu/initiatives/microsoft-365)), which is useful for remote collaboration with other lab group members.

***Regarding Google Sheets***: We understand that some students prefer the collaborative and cloud-based aspects of Google’s applications. Although this week the calculations can be done using Google Sheets, future assignments will be more challenging if you do not use Microsoft Excel. MS Office 365 is still the industry standard (since not all companies will allow their data to be stored on the internet and this is required for Google Sheets), and Office 365 now has collaborative editing features and cloud storage as well. We encourage you to take this opportunity to use the free license and add a line to your resume by becoming proficient in MS Excel, even if you generally prefer Google.

Note: I have a Mac, so the pictures included are from there, but most of this is the same on Windows.

**DOs & DON’Ts:**

* Don’t worry if you know a lot of what is in this assignment, or if you don’t know much of it at all. The point here is to get everyone to a “baseline” level of Excel skill, then, we’ll practice applying these skills for the rest of the semester (and your next physics course).
* Do plan to use the skills you learn in this assignment throughout your physics courses. Keep this document available as a reference!

**Background: *Excel* basics**

This section contains basic information and terminology regarding the use of spreadsheets. It is intended for those students who have little previous experience working with spreadsheets. For those students with a working knowledge of spreadsheets, this should go very quickly.

**Basic terminology:**

1. A spreadsheet is one file in Excel, Google Sheets, Numbers, or a similar program
2. A tab (also known as a worksheet): One page in a spreadsheet. You can add tabs in Excel by clicking on the + sign near the lower left corner. You can rename a tab by double clicking on its title
3. A **column** is a complete vertical set of cells. It is designated by a letter.
4. A **row** is a complete horizontal set of cells. It is designated by a number.
5. A **cell** is one item in the spreadsheet. It may contain text, a number, or a ***function*.** It is identified by a letter/number pair (e.g., A2 or Q38), indicating the column and row.
6. A **range** is a set of cells that may be horizontal, vertical or rectangular. It is designated by the upper left and lower right cells in it, separated by a colon (:). Examples could be part of a column (A3:A23), part of a row (F3:P3), or a rectangle (B2:V22)
7. A **function** is literally a mathematical function. It may contain ***numbers***, ***operators***, ***basic******functions***, and ***references***. **Note**: a function must always start with an equals sign. =(A1+A2+A3) is a function that adds the numbers in three cells.
8. An **operator** is one of the familiar mathematical operators: +, –, \*, /, and ^.
9. A **basic function** is also from math, examples are sum(), exp(), sin(), atan(), and sqrt(). *Excel* has MANY available (modified Bessel functions, anyone?)
10. A **reference** is the value in another cell, named by its column and row as a pair, e.g., A1 or D3.

**References will often serve as the “variable” in our functions.**

1. Some useful syntax is shown in Table 1:

|  |  |  |
| --- | --- | --- |
| Function | ***Excel*** **notation** | **Syntax** |
| π | PI() | π/2 => PI()/2 |
| exponential | EXP(x) | ex  = EXP(x) |
| Scientific notation | 3E-3 | 3x10-3 => 3E-3 |
| Numbers or expressions as text | Enclose in quotes | “sin(PI()/2)+6” |

*Table 1. Useful syntax for Excel Spreadsheets*

To enter information in a cell just click in it and start typing. What you type will appear in the cell and the formula bar (near the top of the window, labeled “***fx****”*). If you need to edit what you have typed, you can work in the formula bar, or double click in the cell. If you type a function, the function will be evaluated when you hit return (or arrow to a new cell).

Now, let’s look at a bit of a spreadsheet. Figure 1, below, was produced on a Mac (since I have one), but the appearance on windows will not be too different.

In Figure 1, cell A2 has the word “Text” typed into it. Cells B2 through G2 have various numbers and functions typed in quotes, so they are not evaluated. Cells B3 through G3 show how B2 through G2 are evaluated without the quotes. Note that cells E3 and F3 refer to cells B3 and C3. Also note that cell E3 has been selected by clicking in it, so it appears in the “function bar” above. Also note that ***Excel* always interprets the argument of sin, cos, etc. in radians, not degrees**. If you want to use degrees, you have two choices: sin(radians(x)) or sin(x\*pi()/180)



*Figure 1. Example Spreadsheet*

**Activity 1 – Complete these exercises on one Excel tab:**

* Exercise 1: Reproduce the figure above
* Exercise 2: Use *Excel* to calculate 4 to the third power. You should be able to get 64!
* Exercise 3: Use *Excel* to calculate the area of a circle with a radius of 10 cm, and to calculate the sin of 60 degrees (which should turn out to be $\frac{\sqrt{3}}{2}≈0.866$).

**Background: Graphs in Excel**

**Scatterplots**

Most graphs used in science and engineering are "scatterplots.” This means that the *x*-axis is defined by one variable (the independent variable) and the *y*-axis is a second variable (the dependent variable). This is very different from simply graphing the *y*-axis data (what *Excel* calls a "line graph"). Figure 2 is an example of this. The data has *x* = {1, 2, 3, 5, 7, 20} and *y* = 3*x* – 2. The graph on the left was made by selecting both the *x* and *y* data and choosing the “X Y (scatter)” chart type in *Excel*. The graph on the right was made by selecting only the *y* data and the line graph chart type. The difference is pretty clear. The scatterplot shows that there is a linear relationship between *x* and *y*. The graph on the right ignores the *x* values and seems to suggest that the *y* data is a curve of some sort.

**When you make a graph in physics lab, you will almost always do so by selecting both *x* and *y* data and choosing a scatterplot from the tool bar or the "insert" menu.**



*Table 2. Example Data*

*Figure 2. (Left) Example scatter plot using the data from Table 2. (Right) Example line graph using only data from the second column of Table 2.*

**Activity 2 – Complete this exercise on a second Excel tab:**

* Exercise 4: Create some data of your own (just make up some numbers to practice with!) and use *Excel* to make a graph of that data. Make sure to label your plot axes!

**Background: Fills and constants**

*Excel* can also be used to evaluate a function for *many* values of the variable. This can be very useful when you are trying to figure out what a complicated function looks like (can you picture sin(*x*)/*x* ? it’s an important function in physics). To do this efficiently, you need to learn **to “fill” a column**, and **to use constants.**

Let’s evaluate the function *ax2 + bx + c* for *x* = [–2, 10] in steps of 0.75, and with *a = 0.2*, *b = –1.5*, *c = 2* as shown in Figure 3 below*.* To do this, I typed the information in rows 1 through 6. The rest of the data was completed automatically by a “fill” operation. You can fill a column (or row) with new versions of a function by either selecting a cell and dragging the blue box in the lower right corner, or by selecting range with a function at the top (bottom or sides also work) and choosing “Fill” (from the edit menu on a Mac, or the home tab on Windows).

Here is how the example works:

Cell A7 contains “=A6+$E$2” (no quotes). When I fill down, *Excel* increments the reference to column A (A6🡪A7, then A7🡪A8, etc.) but it does not increment E2 because **the $ signs indicate “do not increment.”** The result is that each cell contains the value above it plus the constant 0.75. If I showed the function in cell A11, it would be =A10+$E$2.

****

*Figure 3. Example spreadsheet created using the “fill” option for a column.*

Similarly, cell B7 contains the function shown in the function bar. When I fill down, *Excel* increments the reference to column A but does not increment the constants. What do you think it would say in cell B17?[[1]](#footnote-1)

Once the data was there, I used the “X Y (scatter)” choice from the “charts” tab to make the graph (on Windows, it just says Scatter).

**Activity 3 – Complete this exercise on a third Excel tab:**

* Exercise 5: Create a spreadsheet in which column A has rows representing time from 0 to 10 seconds in steps of 0.05 seconds. The top should look something like the figure to the right, and it should go to row 205. You would *not* want to type this by hand! Note: the formula in the highlighted cell in this example says =A11+$A$2. Now, create a function in column B that calculates the function (1/2)*gt*2, where *g* = 9.8 m/s2 is included as a constant in cell B2, with a label above it.
* Exercise 6: Make a graph of your results from Exercise 5.

*Figure 4. Spreadsheet to be*

*created for Activity 3 Exercise 5*

**Background: Basic data analysis**

**Average**

Probably the most basic analytical tool is the average, also known as the “mean” or “mean value.” We (and most scientists and engineers) use the words “mean” and “average” interchangeably. You’ve certainly seen this idea before. If you have 4 values, *x1* through *x4*, the mean of these values is shown in Equation (1):

 $\overbar{x}=\frac{x\_{1}+x\_{2}+x\_{3}+x\_{4}}{4}$ (1)

The variable $\overbar{x}$ is a common way of denoting the average value of *x*. Other common choices are *<x>* and m. *Excel* has a function for doing this: AVERAGE(A1:A10) gives the average of the values in that range.

**Standard Deviation**

Some of you will have seen this before, some not. **The standard deviation is a measure of how “spread out” a bunch of values are.** If you have *N* values, *x1* through *xN*, the standard deviation is shown in Equation (2):

$s=\sqrt{\frac{\left(x\_{1}-\overbar{x}\right)^{2}+\left(x\_{2}-\overbar{x}\right)^{2}+\left(x\_{3}-\overbar{x}\right)^{2}+…+\left(x\_{N}-\overbar{x}\right)^{2}}{N-1}}$ (2)

where $\overbar{x}$ is the average value for that variable. The variable *s* is pretty universally used for this purpose, but *SD* is sometimes used. *Excel* also has a function for doing this STDEV(range).[[2]](#footnote-2) It is important to remember that the spread in data can have several causes. One cause is “noise” in your instrument. That is, you put the same weight on a scale several times, and it gives slightly different readings. Another could be that what you are measuring has an inherent spread. If you want to know the average height of students in a class, you can measure them with a ruler, but no matter how good the ruler is, they don’t all have the same height.

**Standard Error of the Mean**

My guess is that very few students will have seen this before, but it is very important! *SE* (no Greek characters) represents the *uncertainty in the measurement of the average*. The formula is shown in Equation (3), and is simply

$SE=s/\sqrt{N}$ (3)

where *s* is the standard deviation and *N* is the number of samples. It is important not to confuse these last two.[[3]](#footnote-3) The simple thing to remember is when you have a distribution of values, making more measurements doesn’t change *s*, but *SE* keeps getting smaller.

**Example**: Think about the average height of men in Indiana. You can’t measure them all, but you can get a bunch (a sample) and measure them. No matter how many you measure, there will always be some variation in their heights, because there is a built-in spread in how tall people are. That “built-in spread” is measured by the standard deviation. On the other hand, if you measure 500 men, you’ll get a much better estimate of the true average than if you only measure 5. The standard error of the mean gets smaller with increasing sample size.

**Activity 4 – Complete this exercise on a fourth Excel tab:**

* Exercise 7: Here are some numbers: type them into *Excel*, and calculate $\overbar{x}$, *s*, and *SE*.
91, 60, 41, 17, 48, 86, 48, 29, 68, 64, 90, 28, 88, 97, 73, 15, 83, 26, 63, 12

**Percent Error**

Sometimes (not always) you will have a “standard value” for the quantity you are measuring. If that is the case, you can calculate the percent difference between your measured value and the standard value. Note that in research this isn’t the usual case. If there is a standard value, why are you measuring that? There are times it does come up, though, for instance, if you have designed a new measurement system, and wish to show how it compares to previous methods. In this case, the formula is shown in Equation (4):

$\%\_{err}=\frac{|x\_{std}-\overbar{x|}}{x\_{std}} × 100\%$ (4)

where $x\_{std}$ is the accepted (or theoretical) value and $\overbar{x}$ is the measured value. Note that frequently the mean of all measurements is used as the measured value to determine the percent error.

**Background: Curve fits**

Another crucial analytical tool that you can use in *Excel* is fitting data to a function (a line is common, but not required). **The idea is to identify the function that best represents the data.** Once you have a scatterplot of one variable vs. another, you can use *Excel* to determine the line or other function that best represents that data. If you right click on the data in a graph (ctrl-click on a Mac), one of the options that comes up is “Add Trendline” as shown in Figure 4 (below). This will bring up a menu that allows you to choose the function you want to fit, and a place where you can check “Display Equation” This will show you the parameters of the best fit. A linear fit will have a best slope and intercept. A polynomial fit will have the coefficients of the constant, linear, quadratic and higher terms.

Once the trendline is calculated, Excel finds that the data show has a slope of 2.93 and an intercept of 0.27. That is, the function that best represents the data points (0,0.3); (2, 5.7); (3, 8.3)… (10, 26.7) is *y* = 2.93*x* + 0.29. You can try it with those numbers to make sure you are getting the same result.

**

*Figure 4. How to add trendlines (curve fits) in Excel.*

**What data to graph and fit?**

This is another critical skill. Often, your theory tells you what function you expect to describe the data. A simple example is the distance an object has fallen *y* vs. time *t* due to gravitational acceleration *g*: $y=\frac{1}{2}gt^{2}$. If you measure data for *y* vs. *t*, you would not just plot it that way. Instead, you would plot *y* vs. *t2*. That way, the slope of the line you fit is *g*/2!

**Activity 5 – Complete this exercise on a fifth Excel tab:**

* Exercise 8: In your spreadsheet, enter the data in Table 3, make graphs of both y1 and y2, and find the functions that represent the data.



*Table 3. Data for Activity 5.*

* Exercise 9: The data in Table 4 (below) is from a “falling experiment.” Enter the data in your spreadsheet, make the graph suggested above, do the fit, and find *g* from the result. Show the value for *g* in your spreadsheet.

|  |  |
| --- | --- |
| **t (sec)** | **y (m)** |
| 0.5 | 1.27 |
| 1 | 5.07 |
| 2 | 20.04 |
| 3 | 45.82 |
| 5 | 130.20 |
| 7 | 247.83 |
| 10 | 524.07 |
| 15 | 1155.96 |

*Table 4. Data from a “falling experiment” for use in Exercise 9.*

1. I hope you thought it would be =$B$1\*A17^2+$B$2\*A17+$B$3 [↑](#footnote-ref-1)
2. Technically, this is the “sample standard deviation. It is similar, but not identical, to the “population standard deviation.” You can look up the difference if you like, but it will not be crucial here. [↑](#footnote-ref-2)
3. D. G. Altman and J. M. Bland “Standard deviations and standard errors,” <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1255808/> (2005). [↑](#footnote-ref-3)