Spreadsheets in the Physics Lab

At-Home | One Week | Spreadsheet Submission

Introduction

The goal of this first activity is to develop skills using *Microsoft Excel*. While many of you may know some of these skills already, we will use more of Excel's capabilities than most people know. Throughout this semester in physics, we will utilize Excel to record and analyze data, make graphs, do some calculations, and fit functions to data. It is also worth noting that Excel is a widely used tool in a variety of jobs, so the skills you will learn and practice throughout the semester will be valuable for almost any career. This lab is meant to be done at-home, but please get started as early as possible so that you may ask questions of your TA. Your spreadsheet file will be due one week after completing the lab.

Theory

There is very little "physics theory" to cover regarding Excel, but there are many key terms and syntax that will be beneficial to know going forward.

To cover some basic terminology used in this course:

- **Spreadsheet:** A singular Excel, Goole Sheets, Numbers, or similar file.
- **Tab** (or **worksheet**)**:** A single page in a spreadsheet. You can add tabs in Excel by clicking the "+" near the lower left corner of the page.
- **Cell:** A single item box in a spreadsheet. Cells may contain text, numbers, or functions.
- **Column:** A complete vertical set of cells designated by a letter.
- **Row:** A complete horizontal set of cells designated by a number.
- **Range:** A set of cells that may be horizontal, vertical, or part of a rectangle. For example, the range (A3:A23) would include all cells in the A column from row 3 to row 23.
- **Reference:** The column and row associated with a value. Many functions we will use later on will rely on references.
- **Function:** A mathematical function that can be used to calculate values in a range of cells based on numerical values. A function will <u>always</u> begin with an equal sign. For example, if you want to add together the values in cells C13 and D13, you would begin the function in a new cell and type "=C13+D13". <u>Note</u>: do not include the quotation signs in your function.
- **Operator:** A mathematical argument (+, -, *, /, ^)
- **Basic Functions:** Excel has many functions built into it already. For example, say you wanted to add up all of the values in the range A3:A23. You could type "=A3+A4+A5+…+A23", or you could use the *SUM()* function and simply type "=SUM(A3:A23)". These built-in functions

can save you a lot of time so it's worth looking into. Some common readily available functions are: sum(), exp(), sin(), cos(), tan(), sqrt(). **Note:** Excel *always* interprets trig functions (sin(), cos(), tan(),etc.) in *radians,* not degrees.

Some potentially helpful syntax:

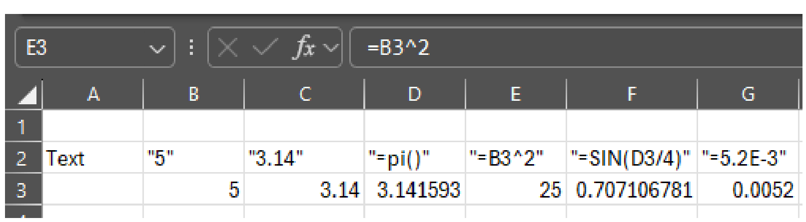
Function	Excel Notation	Syntax
π	PI()	π/2 => =PI()/2
Exponential	EXP()	e^x => =EXP(x)
Scientific Notation	3E-03	$3x10^{-3}$ => 3E-3
Expressions as Text	""	"sin(PI()/2)+6"

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).

Note: *there is an online version of Excel, but while convenient, it has significantly decreased capabilities, so it is recommended not to use the online version for this course.*

Procedure

The following activities are designed to give you some experience using spreadsheets for a variety of tasks. Before we begin, let's see some of the terms defined above in practice. Below is a screenshot from part of a spreadsheet in Excel (I use Windows, so those with Macs may see something slightly different).


E3 | fx | =B3^2

	A	B	C	D	E	F	G
1							
2	Text	"5"	"3.14"	"=pi()"	"=B3^2"	"=SIN(D3/4)"	"=5.2E-3"
3		5	3.14	3.141593	25	0.707106781	0.0052

There are a few things to note about this figure. First, notice how the cells in row 2 are mostly surrounded by quotation marks. This is so that you can visually see the expressions as text. Row 3 is exactly what is in row 2, just without the quotation marks, meaning Excel now "sees" those cells as functions or numbers instead of text and has evaluated those functions. Note how cells E3 and F3 refer to values in other cells (B3 and D3), and note that cell E3 has been selected, so it shows up in the upper left corner and the formula used to evaluate that cell can be seen in the function bar at the top.

Note: *You will be turning in your entire spreadsheet for this assignment, so each activity should be its own tab (labeled by its activity), and each part of the activity should be clearly labeled and easy to find.*

Activity 1: Practice Making a Spreadsheet

1. Reproduce the figure above in your own spreadsheet. Compare the values in row 3 to make sure your spreadsheet is working.

2. In cell H2, write an *expression* for how you would calculate 4 to the third power. Remember, an expression uses quotation marks!

3. In cell H3, write the *function* for finding 4 to the third power. Do you get what you expect ($4^3 = 64$)?

4. In cells I2 and I3 use Excel to calculate the area of a circle with a radius of 10 cm. Cell I2 should have the *expression* while cell I3 should have the *function*.

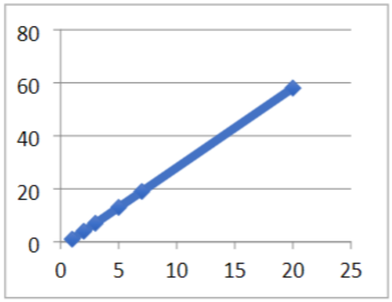
5. In cells J2 and J3, use Excel to calculate the sine of 60 degrees. If you do not get ~0.866, recall the note from the terminology about how Excel handles trig functions like sine, cosine, and tangent.

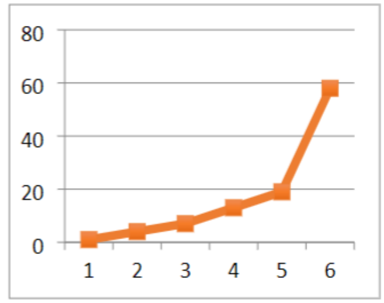
Activity 2: Making Plots in Excel

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"). Here's an example below. The data has $x = \{1, 2, 3, 5, 7, 20\}$ and $y = 3x – 2$. 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.

We will always use scatterplots in physics unless you are told otherwise. So, you will almost always make a scatterplot by selecting both the x and y data and then choosing the "scatterplot" option from the toolbar or the "insert" menu.

x	3x-2
1	1
2	4
3	7
5	13
7	19
20	58





1. Create some data of your own (feel free to make up numbers to practice with). You should have an equal number of *x* and *y* data points and at least 10 rows of data. Use Excel to make a graph of that data.

Activity 3: Using the "Fill" Operation

Excel can also be used to evaluate a function for many different values of a variable. This will be <u>very</u> useful to us in the physics lab, as sometimes we will have to deal with large datasets (looking at you, Conservation of Energy). To do this, you will first need to learn to "fill" a column and use constants. Let's get some practice with this by evaluating the well-known quadratic function ax^2+bx+c for the range of x = [-2 to 10] in steps of 0.75 and setting $a = 0.2$, $b = -1.5$, and $c = 2$.

1. First, put all known constants in the first couple rows of your spreadsheet (how you organize your spreadsheets in the future will be up to you). <u>Make sure to always label each value in your spreadsheet.</u> Your spreadsheet should look something like this:

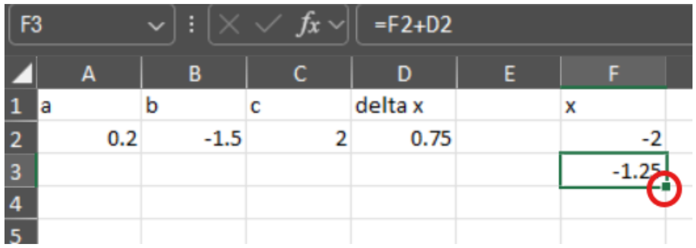
D2 | *fx* | 0.75

	A	B	C	D
1	a	b	c	delta x
2	0.2	-1.5	2	0.75
3				

2. Create a column for *x* in cell F1. Fill in cell F2 with the first *x* value.

3. We need to fill out the F column with all values of *x* in the range *x* = [-2 to 10], while going in steps of 0.75. Cell F2 contains the initial value of -2, so cell F3 should be the result of adding 0.75 to -2. Remember, we already defined the "delta x" constant to be 0.75, so in cell F3 we would naturally want to write: =F2+D2. Try this and verify you get -1.25.

4. Let's now "fill" cell F4 without re-writing the function. To do this, we simply select the cell that contains the function we want to fill (cell F3), then we click on the small box in the lower right corner of that cell (see figure below). We now simply drag that box down once and we should see a number in F4. Drag the function down a few more rows. If you have followed each step exactly as we've outlined, you might see a problem with the values in each cell – they are the same number as the cell above!



5. Before we can fill the rest of column F, we need to tweak our function in cell F3 a bit so that we don't keep getting the same value in each cell. When dragging down formulas to fill new cells, Excel automatically moves reference cells down as well. To see what I mean, select cell F4 and look at the formula. The formula we made referenced cells F2 and D2, however in F4, notice that the formula now references F3 and D3. There's not a value in cell D3, so Excel interprets that as 0 and adds 0 to cell F3, resulting in the same values in cells F3 and F4! To fix this issue, we must tell Excel we want to only reference cell D2 for our step. The way to do this is by using $'s, which tells Excel to stay in only one cell or column. In cell F3, write the expression =F2+$D$2. The first $ tells Excel to stay in column D, the second $ tells Excel to stay in row 2, so together the $'s tell Excel to stay in cell D2. Drag this function down to cell F4 and verify you now get a different value (should be -0.5).

6. Fill the rest of the *x* column up until $x = 10$. Your spreadsheet should end at row 18 and look something like:

F3 | f_x | =F2+$D$2

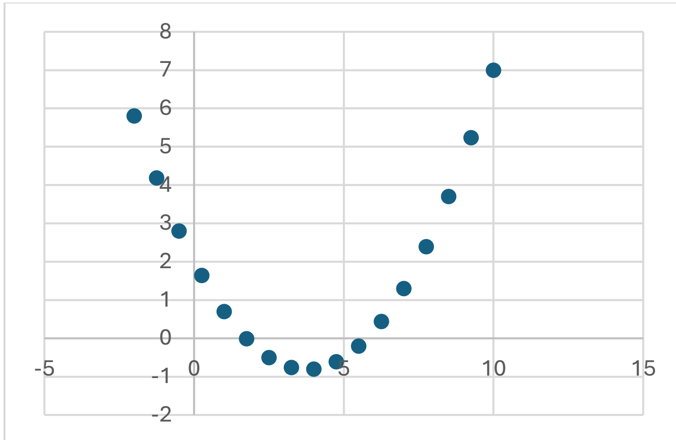
	A	B	C	D	E	F
1	a	b	c	delta x		x
2	0.2	-1.5	2	0.75		-2
3						-1.25
4						-0.5
5						0.25
6						1
7						1.75
8						2.5
9						3.25
10						4
11						4.75
12						5.5
13						6.25
14						7
15						7.75
16						8.5
17						9.25
18						10

7. Create a column for *y* in cell G1 and create a function in G2 for the first *y* value. Fill this column for each value of *x*. <u>Remember to put $'s around all constants!</u> Your spreadsheet should now look like this:

G2 | f_x | =$A$2*F2^2+$B$2*F2+$C$2

	A	B	C	D	E	F	G
1	a	b	c	delta x		x	y = ax^2 + bx +c
2	0.2	-1.5	2	0.75		-2	5.8
3						-1.25	4.1875
4						-0.5	2.8
5						0.25	1.6375
6						1	0.7
7						1.75	-0.0125
8						2.5	-0.5
9						3.25	-0.7625
10						4	-0.8
11						4.75	-0.6125
12						5.5	-0.2
13						6.25	0.4375
14						7	1.3
15						7.75	2.3875
16						8.5	3.7
17						9.25	5.2375
18						10	7

8. Make a scatterplot of this data. It should look like this:



Activity 4: Practice with the "Fill" Operation

Now that we have some background, let's try practicing this with a different data set and function. This time, x and y will represent real physical variables (time and position, respectively). The time will range from 0 to 10 seconds, but let's go with a very small time step of 0.05 seconds. For the position, we will use the function $(1/2)gt^2$, where g will be a constant 9.8 m/s^2.

1. Create your spreadsheet using the given constants and variables. Remember, each value in a spreadsheet should be clearly labeled! I will not show the entire spreadsheet or the function syntax, but the first few rows of your spreadsheet should like this:

	A	B	C	D	E
1	time (s)	y (m)		delta t (s)	g (m/s^2)
2	0.00	0.00		0.05	9.8
3	0.05	0.01			
4	0.10	0.05			
5	0.15	0.11			
6	0.20	0.20			
7	0.25	0.31			

2. Create a graph of this data. Remember, time is the *independent* variable and position is the *dependent* variable.

Activity 5: 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, x_1 through x_4, the mean is $\bar{x} =$

$\frac{x_1+x_2+x_3+x_4}{4}$. $\bar{x}$ 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, x_1 through x_N, the standard deviation is

$$s = \sqrt{\frac{(x_1-\bar{x})^2+(x_2-\bar{x})^2+(x_3-\bar{x})^2+\cdots+(x_N-\bar{x})^2}{N-1}}.$$

s is pretty universally used for this purpose, but *SD* is sometimes used. *Excel* also has a function for doing this STDEV(range)[1] 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) represents the *uncertainty in the measurement of the average*. The formula is simply $SE = s/\sqrt{N}$. It is important not to confuse these last two[2]. 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.

For 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.

1. Given the following values, use Excel to calculate the mean, standard deviation, and standard error: 91, 60, 41, 17, 48, 86, 48, 29, 68, 64, 90, 28, 88, 97, 73, 15, 83, 26, 63, 12

---

[1] 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.

[2] D. G. Altman and J. M. Bland "Standard deviations and standard errors," https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1255808/ (2005).

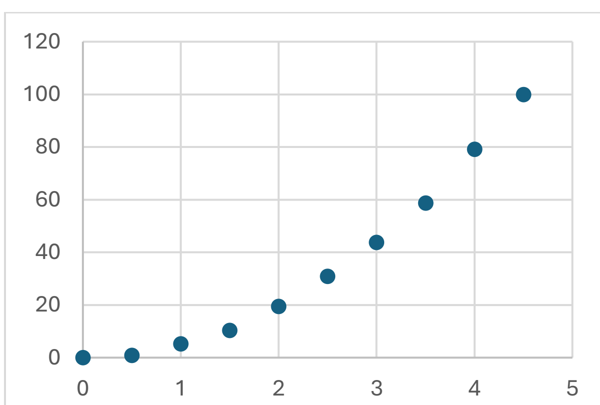
Activity 6: Fitting Functions to Plots

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." 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. Let's try this out with some sample data.

1. In a new spreadsheet, enter the data below:

Time (s)	Position (m)
0	0
0.5	0.9
1	5.3
1.5	10.4
2	19.5
2.5	30.9
3	43.8
3.5	58.7
4	79.1
4.5	100

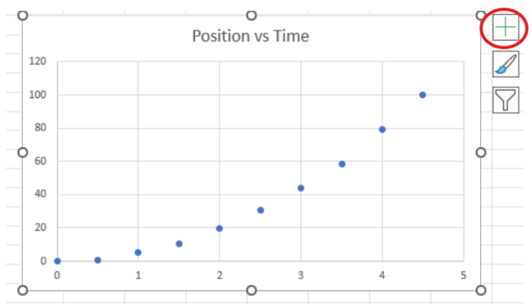
2. Make a graph of this data. It should look like this:



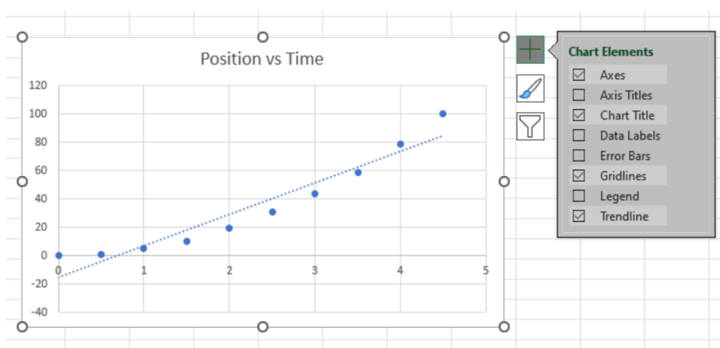
3. Before fitting any trendline through the data, look at the shape of the graph. You can usually "guess" what type of trendline (polynomial, linear, exponential, logarithmic, etc.) you should

start with based on the shape. From the shape of the graph above, we can reasonably rule out linear and logarithmic fits. This leaves us with basically two choices: polynomial and exponential. <u>Polynomial and exponential fits are not the same</u>. They fit two very different lines. Without knowing the theory behind the data, we have no real sense as to which fit would be best, but it's best in this course to simply default to polynomial fits as nothing we will explore in this class requires an exponential fit.

4. Add a polynomial trendline to the graph. To do this, we will use the plus button on the upper right-hand corner of the graph when selecting it in Excel (Note: fitting curves is the primary reason the online version of Excel is not recommended, as it is significantly more challenging to do this in the online version). The plus button is also how you will add things like x-axis and y-axis labels.



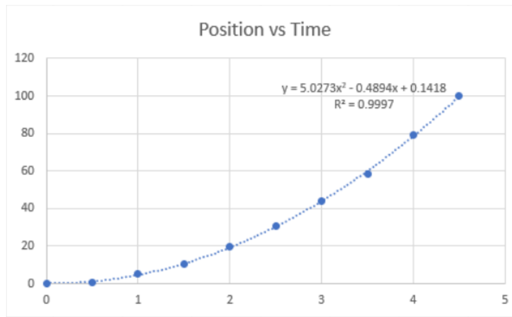
5. Select "Trendline." Note that the default trendline is linear.



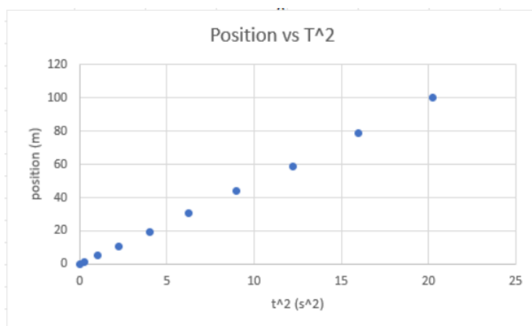
6. Go to more options by selecting the arrow next to trendline. Then select "polynomial."

7. We want to see the equation displayed on the chart, so at the bottom we will select "Display Equation on chart" and "Display R-Squared value on chart." If done correctly, your graph should look like the one below. The equation displayed on the chart is the equation of the line fit. The R-squared value is essentially a measure of how "good" your fit is. The closer the value

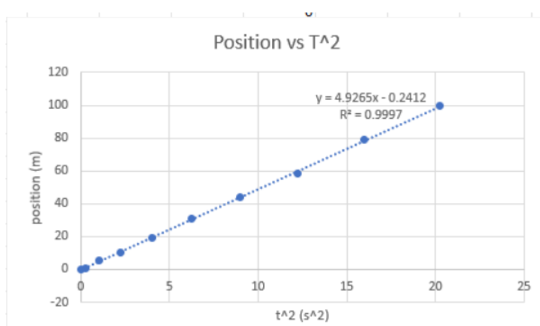
to one, the better the equation fits your data. Try fitting this data with a linear fit and compare the R^2 values to get a better feel for how to interpret this value.



8. Sometimes we can "linearize" data to make it more easily interpretable. This usually requires at least some theoretical background knowledge of the data, but it is not always required. Based on our graph above, it appears that a polynomial fit of power 2 (i.e., a quadratic equation: $ax^2 + bx + c$) fits the data best. So, let's linearize this data. Create a new column in your spreadsheet and call it t^2. Square the time values (you should know how to do this from Activity 3) to fill out this column. Make a new graph of the position vs t^2. It should look something like this:



9. Fit a linear trendline through this data. Remember to display the equation and R-squared value. Your finished product should look like this:



Activity 7: Using Plots to Determine Variables

Finding the right function for a set of variables is only part of the picture for many of the experiments we will do. The other key skill to develop is using those fitted functions to find a value of interest. Let's assume the values in Activity 6 are from a simple falling experiment of a ball being dropped by some initial height. This is very well described using the kinematics equation $y_f = (1/2)gt^2 + v_it + y_i$. We don't need to get into the nitty-gritty details of this equation yet (that's for your lecture), but it's helpful to at least define the terms. The initial and final positions are given by y_i and y_f, respectively. The initial velocity is given by v_i, the acceleration due to gravity is given by g, and the time is represented by t. Let's say our goal is to find g, so we need to use our first graph, and the equation given, to find g.

1. The equation from our first graph is $y = 5.0273x^2 – 0.4894x + 0.1418$. Using the kinematics equation given, let's match up variables.

$$Given: y_f = \left(\frac{1}{2}\right)gt^2 + v_it + y_i$$

$$First\ Graph: y = 5.0273x^2 - 0.4894x + 0.1418$$

   It should be somewhat obvious how to match these values up when looking at them in this way. First, let's start by matching the axes. The final position, y_f, is clearly represented by the y in the equation from the graph. This tells us that each point on our graph represents the *instantaneous* final position at that time. We graphed the time, t, on the x-axis, so we know that the time, t, is represented by the x's in the graph equation. That leaves y_i, v_i, and g left. In the given equation, the last value is y_i and is not accompanied by a t, so we should look for a single value without an x in the graph equation, which is obviously the 0.1418 value, indicating that according to this graph, the initial starting point (y_i) is at 0.1418m. Following this same formula, v_i = -0.4894 (note the negative). Finding g is slightly trickier, since the given equation contains a (1/2). Still though, the same process is applied such that $(1/2)g = 5.0273\ m/s^2$. Meaning that $g = 2*(5.0273\ m/s^2) \approx 10.0546\ m/s^2$. Make a data table to summarize these results.

2. Let's follow the same procedure, but this time using the linearized graph from Activity 6. We took steps to linearize the data, but an important assumption was made when we did this. By linearizing the data, we are assuming that the data should be linear, in which case the initial velocity and initial position would both be zero. Think about why we must make that assumption. If we set v_i and y_i to zero in the given equation above, what does that look like? Well, it would look something like this:

$$Given: y_f = \left(\frac{1}{2}\right) g t^2$$

This looks remarkably similar to the simple equation of a line that we all know and love:

$$Linear\ Equation: y = mx + b$$

So in this case, our slope value should be (1/2)g. Or rather: $g = 2*slope \approx 9.853$ m/s^2. It is worth mentioning that in this case, the x term in the linear equation is represented by t^2, not just t as was the case before.

Again, make a data table to summarize these results.

Activity 8: Practice Fitting Functions & Determining Variables

1. Enter the following data into a new tab for y1 and y2. Graph y1 and y2 (separately). Find the functions that best represent each data set.

x	y1	y2
0	0.4	0.7
1	3.3	3.4
2	6.2	11.6
4	11.1	44.2
6	19.3	114.9
7.3	25.0	176.8
8	28.1	189.4
9.1	32.3	265.7

2. The data set below is to a "falling experiment" in which an object was dropped from rest and the height at various times was measured. Enter the data in your spreadsheet, make a graph of this data, do the fit, and find g from the result.

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

FAQ’s & Recommendations

How should I prepare for lab time?

You only have so much time in lab each week, so proper preparation makes a huge difference in what you’re able to accomplish! Read the handout ahead of time so that you can ask clarifying questions immediately and get started as soon as you arrive!

What goes in my lab notes?

The purpose of lab notes is to enable your or a colleague to reconstruct what was done and why after you’ve left the lab and are performing analysis or writing a submission.

- You can use any form you like to record experiment information: notebook, spreadsheet, etc.
- They don’t have to be neat, in complete sentences, etc., but they do have to be useful!
- Make sure to take detailed notes about your setup, how to use the equipment, what results you found, measurements related to the environment you may need, etc. You may not be able to get back into the lab later in the week if you miss something, so record as much detail as possible!
- When storing multiple data files while in lab, make sure to name the files clearly so they’re easy to find later.

When should I work on the experiment and analysis?

We strongly recommend doing the lab as early in the week as possible, rather than waiting until it is almost due. This is just so that, if you run into trouble and need help, you’ll have plenty of time to talk to your TA and get issues resolved before the deadline.

How do I turn in my results?

After leaving lab, performing your analysis, and completing your submission, you’re ready to turn in your work!

- Every lab session requires submission of either an assignment, summary, draft report, or report.
- Collaborate with your partners on data collection, analysis, and writing.
- Turn in a single group submission and make sure the names of all group members are included.
- Upload your submission to Canvas/Brightspace as a .pdf by the deadline in the course calendar.
- Other than the spreadsheet assignment, you will not upload any spreadsheets. Just copy and paste figures and other elements from your spreadsheet into your formal submission as needed.

Where can I get help?

Your lab TA can answer questions during the lab, by email, or by setting up a time to meet. You can also ask advice from lab partners and/or other students.

General DO’s and DON’T’s

- *DON’T* break the equipment – always be careful when using lab supplies!
- *DO* consult with your lab TA before leaving a lab session about your experimental method, the validity of your results, and any confusion you have about the analysis process.
- *DON’T* forget to record all the parameters and measurements for your experiment, including saving files.
- *DO* be creative in your experimental design and enjoy!