Error Analysis in Physics Lab

In-Person | One Week | Spreadsheet Submission

Introduction

When performing any experiment, many students assume the experiment is done when a measurement is made and recorded. This assumption ignores one of the most important questions any scientist must ask themselves when making *any* measurement: *How confident am I in this measurement?*

Consider two scientists, Kara and Avi measuring *g*, the acceleration due to gravity. Each measures *g* 10 times. Their results are

Kara: 9.738, 9.734, 9.764, 9.761, 9.752, 9.707, 9.837, 9.789, 9.790, 9.716

Avi: 8.635, 12.136, 12.000, 8.593, 12.143, 9.278, 8.816, 8.920, 8.293, 9.756

Who has done a better job? Averaging the numbers, Kara gets a value of 9.759, whereas Avi gets a value of 9.857. Both are very close to 9.81, but Kara's values are close to one another, and Avi's are all over the place. To make this clear, look at what happens if they had each stopped after 8 measurements. Kara's average would be 9.760, but Avi's would be 10.065!

The same thing happens if they each measure a few more times, Kara's average is unlikely to change much, whereas Avi's average could change quite a bit. We will discuss the difference between *accuracy* and *precision* a little later on, but this does a reasonably good job of illustrating why uncertainty matters!

As a scientist, we must report uncertainties whenever possible, but in order to do this, we first must explore the concept of *error*. It is important to note first that error is a natural component to *any* experiment, just as imperfections are a part of *any* manufacturing process. The goal is never to eliminate error, the goal is *know what the errors are*, and to get them to be *small enough*.

This lab is designed to give you practice identifying and quantifying sources of error in experiments. Error analysis is an important component to each lab, and you will need to use the skills you develop today in the remaining labs this semester. Please continue to refer to this handout for error analysis, and taking your own notes is encouraged.

Theory

Sources and Types of Error:

Every experiment contains some degree of error. **Errors and mistakes are not the same thing**.

A mistake is something that you did wrong and can simply redo. For instance, you weighed an object and forgot to subtract the weight of the pan, or you read a meter that said 10.13 Volts and wrote down 1.13 V. **If you make a mistake, you should discard the mistake and take the data over again.**

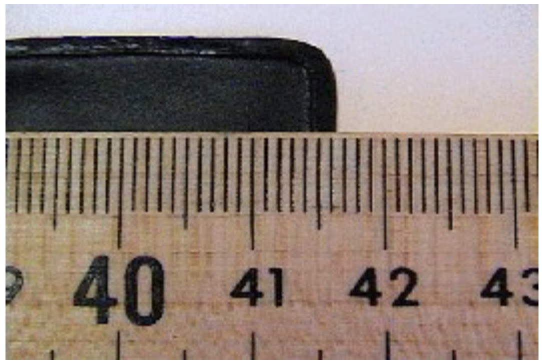
Errors are not mistakes, they come from the equipment you are using. Redoing the experiment won't change anything because the errors are inherent to the measurement. **The way to deal with the errors is to quantify them and report them as an uncertainty in your outcome**. If the uncertainty is too high for your purpose (and if your budget allows it) you can reduce the errors by repairing an instrument (if it is faulty) or replacing it with a different type (that produces smaller errors).

We identify errors of two types: *systematic errors*, and *random errors*. We will frequently make assumptions in labs to make measurements easier (frictionless ramps, anyone?), and often these assumptions will result in sources of error. In fact, these assumptions are an example of a *systematic error*. <u>Systematic errors affect all measurements equally and in the same way.</u> For example, let's say you are asked to measure the length of a steel rod with a meter stick. Unfortunately, the meter stick is slightly warped, so each measurements would be off by the same amount based on the warp of the meter stick. Let's say we use a laser to measure the length of the meter stick we used and find that the warp causes measurements to be 0.200 cm short. In this case, we can actually quantify this source of error as –0.200cm. Because all length measurements would be affected equally, the warp of the meter stick would qualify as a *systematic error*. **Systematic errors are usually predictable but often very hard to identify.** If you didn't have that laser tool, it would be tough to notice this error, and impossible to say exactly what it is.

By contrast, there are other sources of error that are unpredictable. We call these unpredictable sources of error *random error*. This is the type of error that is often called *noise*. <u>Random errors affect each measurement individually.</u> Let's go back to our example of length measurements. Now, we measure the rod 5 times using the laser and get the values 13.00cm, 12.74cm, 13.12cm, 13.03cm, and 12.97cm. The variance in these repeated measurements is due to *random error*. **Random errors are typically harder to quantify than systematic errors due to their unpredictability, but they still can be quantified**. We will explore this further but consider what could be the source of error that would have caused the variance in the string length measurements.

An experiment can have any number of sources of error, and it's important that we can describe them. **Sources of error lead to uncertainties in a measurement.** An important component to every lab will be a discussion on sources of error. You will be asked to describe what errors could affect the results of your experiments, and it is important to consider both *random* and *systematic* errors. For example, let's once more go back to our experiment. You measure the same rod 5 times and get 5 different answers. We know that each answer will be off from the "true" value because we know the meter stick was warped. The warp of the meter stick is a source of systematic error. But where else does error come from? You might be tempted to use the term "human error," but that implies "mistakes." It is too broad of a term to be meaningful in error discussion because it can't be quantified. Instead, let's think about the experiment itself. You were asked to measure a rod with a meter stick. We already discussed how the meter stick is warped, but there is another component that we haven't mentioned yet, and that's in the measurement itself. Consider the following diagram[1]:

[1] https://science.clemson.edu/physics/labs/tutorials/measure/index.html



In the diagram, note the tick marks of the meter stick and where the object falls. The end of the object falls *between* the tick marks. That means when we measure the length of the object, there is a certain amount of *estimation* that has to happen. I might read the length as 41.63cm, while you might read it as 41.65cm. That last digit has to be estimated, which leads to error. This doesn't mean you have made a mistake in measuring the object, it means you have to use your best judgement. The *estimation* is the source of error because you have to use judgement and approximation. So while the term "human error" is too broad, *estimation based on the resolution of the meter stick* is not.

Procedure

For this lab, we will utilize the following equipment:

- Meter Stick
- Tennis Ball
- Stopwatch (x2 for groups of 3)
- Excel (or equivalent) for data analysis and assignment preparation

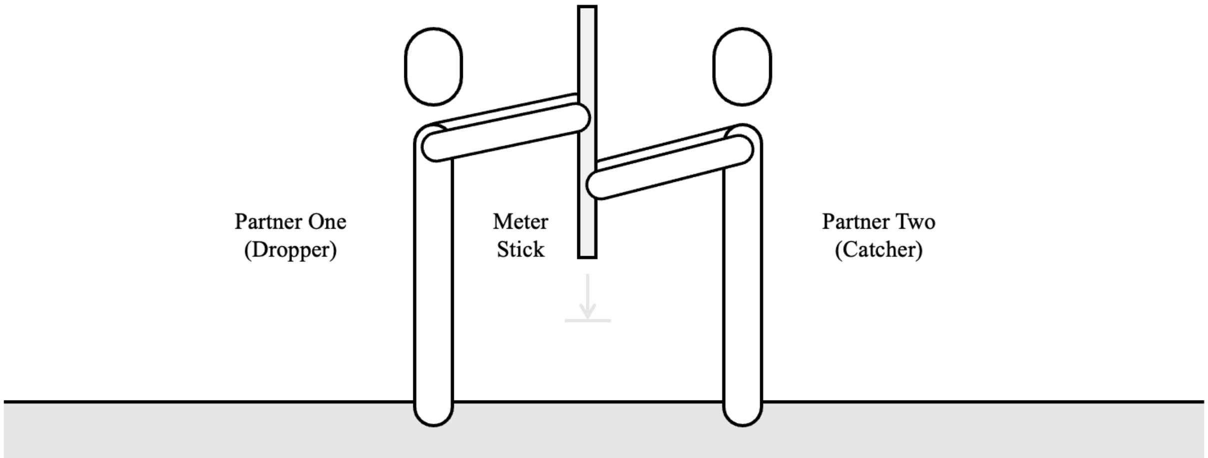
We will explore quantifying and measuring sources of error using real experimental data. Activities 1 and 2 should be done in separate spreadsheet tabs (see the *Spreadsheets* lab for common terminology if not familiar with Excel (or equivalent). Your spreadsheet will be due one week from the end of this lab.

For both activities, we will be dropping an object from rest. Without getting too into the weeds (that's what lecture will soon be for), we can describe this motion using the equation:

$$\Delta y = \left(\frac{1}{2}\right) g t^2,$$

where Δy is the change in the object's position, g is the acceleration due to gravity (≈ 9.80 m/s^2), and t is the fall time of the object.

Activity 1: Measuring Reaction Time



1. Choose one lab partner to be the "dropper" and one to be the "catcher." Each partner will take a turn in each role, so it doesn't really matter which you begin with.

2. Partner 1 (the "dropper") should hold the meter stick near the top (see experimental diagrams). Partner 2 (the "catcher") should hold their hand near the midpoint of the meter stick. Make sure the catcher is not touching the meter stick but is ready to catch it.

3. Record the initial position of Partner 2's hand. Keep track of where on Partner 2's hand you are making the measurement from.

4. **Without warning**, Partner 1 will drop the meter stick. Partner 2 will then catch it as fast as possible and hold the meter stick. If the meter stick falls through without being grabbed, you may consider beginning slightly lower on the meter stick.

5. Record the final position of Partner 2's hand. From the initial and final positions, determine how far the meter stick fell before being caught.

6. Repeat steps 2-5 for a total of at least 5 trials while keeping the same partners as the dropper/catcher. A sample *Excel* data table for this data might look something like this:

Partner	N	Avg. t (s)	s (s)	SE (s)
1				
2				
3				

Trial	initial height (m)	final height (m)	change in height (m)
1			
2			
3			
4			
5			

7. For each trial, use the equation above and the change in height to find the reaction time.

8. Find the overall **average** reaction time for Partner 2. We use the terms "average" and "mean" interchangeably. If you have 5 values, x_1 through x_5, the mean (or average) is $\bar{x} = \frac{x_1+x_2+x_3+x_4+x_5}{5}$. $\bar{x}$ is a common way of denoting the average value of x. *Excel* has a function for doing this: for example, =AVERAGE(A1:A10) gives the average of the values in the range from cell A1 to A10.. Record (and label!) this value in your spreadsheet.

9. Find the **standard deviation** of the reaction time values. <u>*The **standard deviation** is essentially a measure of how "spread out" a bunch of values are.*</u> you have N values, x_1 through x_N, the standard deviation is

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

   The symbol s is pretty universally used for this purpose, but *SD* is sometimes used. *Excel* also has a function for doing this STDEV(range). Record (and label!) this value in your spreadsheet.

10. Find the **standard error** of the overall average reaction time. <u>*The **standard error** represents the **uncertainty** in an average.*</u> It is very simple to calculate once you know the standard deviation, as it is scaled by the number of values, N. The formula for finding the standard error is:

$$SE = \frac{s}{\sqrt{N}}$$

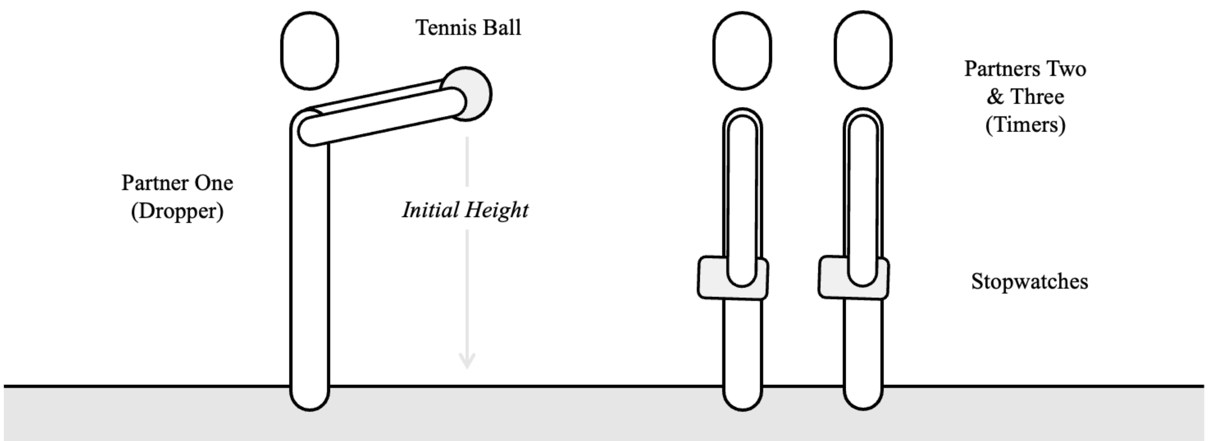
   Remember, s is the standard deviation and N should be the same value in both the average and standard error calculations. Record (and label) this value in your spreadsheet.

11. Repeat steps 2-10 for each partner as the catcher.

12. Create a data table to summarize each partner's average reaction time and uncertainty. It might look something like this

13. Come up with 3 possible sources of error and list them in your spreadsheet. Additionally, briefly discuss how each would affect your measurement (like how the warped meter stick in the string example would cause our measurements to be off by ~2cm). You do not need to put a numerical value with each one, but at least one source of error should be quantified (i.e., with a numerical +/-).

Activity 2: Presenting Errors in Reports



Another component to performing an error analysis is knowing how to present the error in a report. Sources of error are important, but we still must answer the question: *How well do I know this value?* There are multiple ways we can do this. One we've already discussed, which is the standard error. **Every final result reported should have an associated uncertainty.** For example, if we were asked to report the string length for the first string in our string length example and we were to report just the average of the values (12.97cm), this would not be complete. We would instead report 12.97 +/- 0.06cm, where the 0.06cm is the standard error and indicates the uncertainty in our result. The uncertainty in a measurement is usually denoted with either a Greek delta (Δ) or sigma (σ). To avoid confusion with delta also meaning "change," we will use σ to denote uncertainty (e.g., the uncertainty in g would be written as σ_g).

The same is true for graphs. Often in physics, you will be asked to find a value based on a graph (in fact, some of you will have already done this from the *Excel* lab), and we still need a way to find the error in these values. Thankfully, we can leverage Excel's LINEST function to do this for us. We will go into depth about this function in a little bit, but for now just know that it uses the "least squares" method to calculate the statistics of a straight line[2].

---

[2] For more information, see: https://support.microsoft.com/en-us/office/linest-function-84d7d0d9-6e50-4101-977a-fa7abf772b6d

This activity is designed to give you experience and practice presenting errors in graphs and presenting/interpreting results. You will drop a ball from various known heights and record the time it takes for the ball to fall. From there, you will make a graph and find the acceleration due to gravity, *g*.

1. Pick a starting height somewhere around the room. Feel free to spread out into the hallway as necessary but remember that classes may be going on around you, so please keep your noise to a minimum. Measure and record this height.

2. One partner will be responsible for dropping the ball, while the other two will be responsible for timing the fall.

3. Drop the ball and measure the time it takes to fall. If you are unable to react quickly enough, consider increasing the height of the drop. Record both times in a spreadsheet.

4. Repeat for a second trial. Your spreadsheet may look something like this:

trial	t1 (s)	t2 (2)
1		
2		

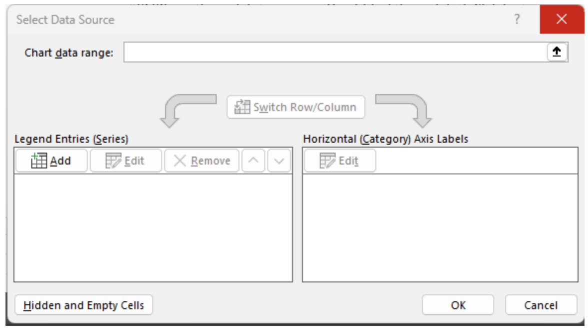
5. Average all 4 times together and find the standard deviation and standard error.

6. Repeat steps 1-5 for at least a total of 4 different heights.

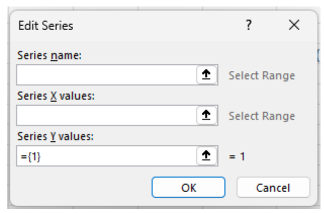
7. Make a summary data table of the results. It should look something like this:

Height (m)	t_avg (s)	t_SE (s)
1.00		
1.25		
1.50		
1.75		
2.00		

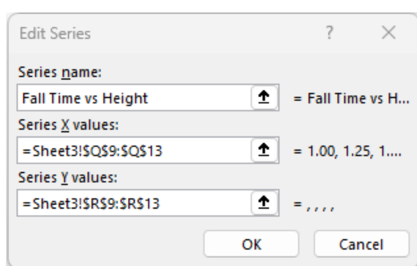
8. Make a scatterplot of the data. In Excel, this would mean selecting "Scatter" from the list of charts in the "Insert" tab from the ribbon. From here, you will right-click the graph and select "select data." You should see the following pop-up:



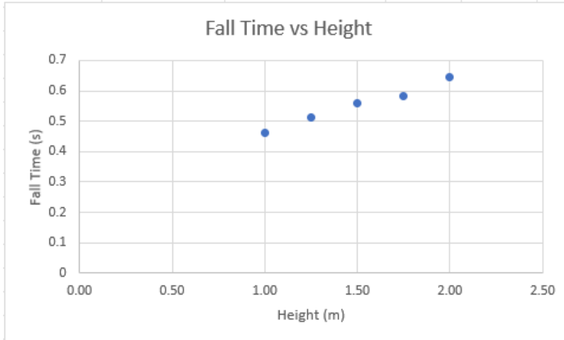
9. From here, you will select "Add" under the Legend Entries on the left side. You should see the following:



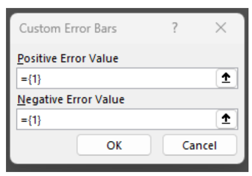
10. You can give the series any name you want. Conveniently this will also name the graph, so something like "drop time" or "drop time vs height" would be fine. The other two lines are where you will select your data. In the "series X values" box, type an =, then go to your summary table where you have your height values listed in order, and highlight them. Delete the ={1} under the series Y values and do the same for the average times. My heights are in cells Q9 through Q13 and my times are in R9 through R13, so my filled out series looks like:



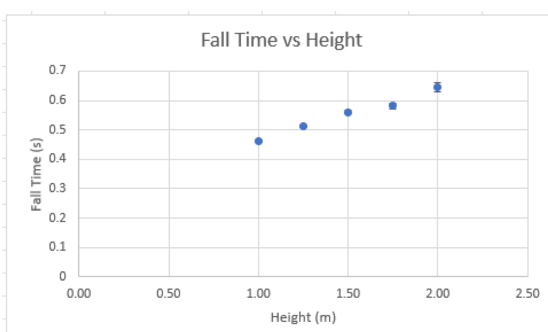
11. Click OK. You should see a graph that looks pretty linear. You should add axes labels using the + sign in the top right corner of the graph once you select it. Every graph you will submit should have axes labels *and* units. Up to this point, your graph should look similar to mine:



12. Now, we want to convey to the readers the relative size of the error for each data point. To do that, we need to add error bars to the graph, and we need those error bars to represent the standard errors we found earlier. From the + menu, select "error bars" and just like for the trendline, go to "more options." We will never use horizontal error bars, so you can simply delete those. Excel has options for error bars by default, but Excel doesn't know what it doesn't know. You will see an option for standard error under Error Amount, but Excel doesn't know what values you used to find the standard error, so that option won't result in what we want. Instead, we need to select the "Custom" error bars. You should see this pop up:



13. Delete the information in each box already, start with = and then highlight the cells which include your standard error. Mine are in cells S9 through S13. You will use the same values both positive and negative error bars. Your graph should look very similar to the one below. Notice how only the last two data points have error bars (and notice how small they are). That's ok, it just means that relative to the size of the plot, the error bars for the first three data points are too small to be shown.

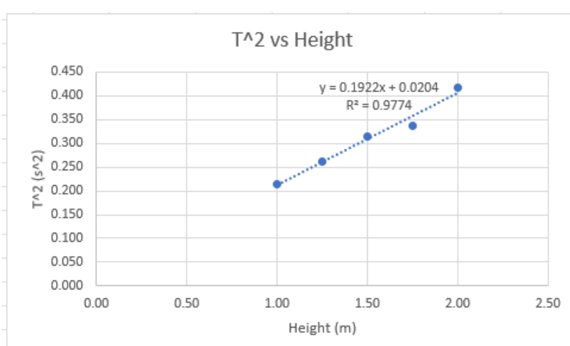


14. Now, make a new column in your summary table and square all of your time values. Mine looks like this:

Height (m)	t_avg (s)	t_SE (s)	t^2 (s^2)
1.00	0.462	0.005	0.213
1.25	0.512	0.003	0.262
1.50	0.561	0.007	0.315
1.75	0.581	0.010	0.338
2.00	0.645	0.015	0.416

15. Make a new graph of t^2 vs height. That means t^2 will be on the y-axis and the height will be on the x-axis.

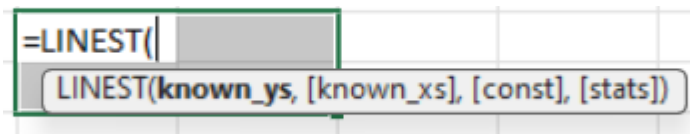
16. Let's practice adding a trendline to this new graph. The +-sign we used earlier to add the axes labels can also be used to add trendlines, legends, and error bars. Practice adding a trendline to the graph. Clicking on the arrow right next to trendline includes, "more options." From here, you can change the function that the trendline uses to fit (though we'll keep it on linear), and you can change other options like forcing the trendline through zero, displaying the equation, and displaying the r-squared value. Select the options to display the equation and r-squared value. **The r-squared value is essentially a numerical representation of how well the function fits the data.** An r-squared value of 1.000 indicates a perfect fit with zero deviation (and should be immediately suspicious for an experimenter). A value of 0 would indicate no correlation at all. Your new graph should look similar to this (note the change in the y-axis labels):



17. Without getting into the details too much, the slope of this graph is very interesting. Based on the equation $y = (1/2)gt^2$, we can rearrange it slightly to give $t^2 = (2/g)y$. If we compare this equation to that of a line : $y = mx+b$, the position is given by x in the line equation, t^2 is represented by y, which leads to the slope, m, represented by 2/g. Find g from your slope. Record this value.

18. Now, as we said before, we must always give an uncertainty with every reported measurement. To find the uncertainty in our slope, we use the LINEST function built into Excel. It's not overly important to know the details, but if interested, you are welcome to see the resource listed at the bottom of page 4. To begin, select a 2x2 array of empty cells in Excel.

19. Type =LINEST. You can double-click the predictive option Excel gives you, or you can simply open a parenthesis. Either way, it should look like this:



20. Then we simply fill out each part. Highlight your known y's (the t^2 values), then add a comma and highlight your known x's (the heights), then add another comma and type TRUE, then add another comma and type TRUE again. **Don't just hit enter!** You need to press shift+control+enter. Otherwise, you will only get one value in the array. Your result should be a 2x2 array of numerical values. I've added borders and labels around mine to illustrate what each component represents:

Slope	0.192232	0.020432	**y-int**
unc. In slope	0.016866	0.025993	**unc. In y-int**

21. The top and bottom left values are the ones we will use the most. The top-left value is the slope and the bottom-left value is the uncertainty in the slope, calculated using the least-squares method. The upper-right value is the y-intercept, and the lower-right value is the uncertainty in the y-intercept. Make sure you can get this 2x2 array. If not, make sure to ask your TA for help.

22. Earlier, we found g by using g = 2/slope. The uncertainty in g can be found similarly as $\sigma_g = 2/\sigma_{slope}$. (This method is admittedly a little hand-wavy, but it does a reasonable job at this level). Find the uncertainty in your g-value. Record this value in your spreadsheet.

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!