Lab 0: Data Sheet Guide

Aaron Dunbrack

Summer 2018

Contents

Regardless of how comfortable you are with Google Sheets, you should read sections 1.3 (which goes over about how to use the `LINEST` function) and 3 (which talks about the specifics of how to use Google Sheets for this course).

If you don't know how to do use your spreadsheet program as a really powerful calculator, then it's worth reading through the rest of section 1 as well.

For the most part, the plotting tool should be straightforward using the manual interface, but if you want to make your life a little easier, read section 2.2.

1 Calculations in Google Sheets

If you've only used spreadsheet programs[1] occasionally, you may have only used them to put numbers in some organized form.

If so, that's a shame. These programs are very powerful, and learning even the basics of how to use them greatly expands your ability to efficiently analyze quantitative data.

Here, instructions will refer to Google Sheets, because that is a program that everyone with an internet connection has access to. The general abilities discussed here, should apply to virtually any spreadsheet program, however.

1.1 Using Formulas

1.1.1 Basic Calculations

The first thing one should understand about Google Sheets is that it has a built-in calculator.

Whenever you start a cell with an equals sign, Sheets treats everything that follows as a calculation.[2] For instance, if you enter `=2+3` in a particular cell, that cell will appear as `5`. (Try it!)

If you double-click on the cell, it will return to displaying `=2+3`. Thus, you can enter a formula into a cell, then go back and edit it later. You can also look at the text bar at the top of the screen, where it will display the full formula; clicking there will also enable you to edit the formula.

1.1.2 Functions

In addition to the usual operations (`+`,`-`,`*`,`/`,`^`), Google sheets has a number of functions you can use. For instance, if you want to calculate the square root of 2, you can write `=sqrt(2)`.

Some that you may find useful:

- `sqrt(x)`: As mentioned.
- `pi()`: If you want to insert π, rather than typing 3.14, you can use this function. (Remember the parentheses - it won't work without them.)

[1] Excel, Google Sheets, OpenOffice Calc, LibreOffice Calc...

[2] If you ever want it *not* to do this - say, you want a cell that actually starts with an equals sign - put a single quote, ', before the equals sign.[3] This forces everything after to be treated as ordinary text rather than a calculation.

[3] If you want a cell to actually start with a single quote for some reason, start with two single quotes instead. The first one will "mark it as text," and the second one will appear.

- `sin(x)`, `cos(x)`, etc.: Note that these take arguments in *radians*, so if you have your numbers in degrees you'll need to convert them. Mulyiply your angles by `pi()/180` to do this conversion, of course.

- `asin(x)`, etc.: this is how you write the function also called arcsin or $\sin^{-1}$. Note that these return angles in radians, too.

- `exp(x)`: $\exp(x)$ means the same thing as e^x.

- `ln(x)`, `log10(x)`, `log(x,b)`: logarithms in various bases. In this class, we'll only use `ln`.

There's also a number of more mathematically advanced functions included built-in, in addition to some general analysis tools. A more complete list can be found at `https://support.google.com/docs/table/25273?hl=en`.

1.1.3 Calculations Using Cells

The real power of Google Sheets comes from the fact that you can do calculations not only with numbers, but with *cells*.

For instance, you can have an expression like `=A1+A2`. If you enter this into, say, cell `A3`, then cell `A3` will automatically be filled with the sum of whatever is in cell `A1` and cell `A2`. (Empty cells are treated as zero.)

This means that you don't have to re-enter numbers. You can just refer to the cells those numbers are in. This lets you do calculations directly in Excel very easily.

That's kinda convenient, but it actually gets better. Whenever you change a cell, Google Sheets *automatically* updates anything that's based on that cell.

For example: suppose you started with a `1` in cell `A1` and a `2` in cell `A2`. If cell `A3` was `=A1+A2`, then it would show `3`, of course. If you then changed cell `A1` to a `5`, then cell `A3` would automatically change to a `7`.

1.1.4 Usefulness of Cell Calculations

Alright, it's neat that the spreadsheet works that way, but how would this help you in lab?

Well, let's suppose you have a sequence of calculations across multiple cells. If you do something wrong in the middle, you can fix that calculation, and it'll automatically fix the last one!

For instance, suppose cell `A3` is `=A1+A2`, and cell `A4` is `1/A3`. You then realize that, oops, you made a mistake, and `A3` should be `=A1-A2` instead.

If you had done this in the calculator, you would have to recalculate both. In Google Sheets, you can change the formula for `A3`, and `A4` will change automatically. Nice and convenient!

Actually, if you're ambitious, this can help even more. If you read the lab beforehand and understand what calculations you're going to be doing, you can do all your calculations ahead of time!

Before you get to class, enter all the formulas into the data sheet. Then, when you get to class, you can enter in the data, and your final results will come out as you take your data. This means you can check whether your results are reasonable before even leaving the lab, without having to do a bunch of work during lab to get that done.

1.2 Using Google Sheets Autofill

1.2.1 Pattern Recognition

The second trick that makes spreadsheet tools incredibly useful is their ability to pick up on patterns.

Let's suppose you want your first column to have the numbers 1-100. If you enter the first few numbers into the first few cells, you can have Google Sheets do the rest.

Enter `1` into `A1`, `2` into `A2`, and `3` into `A3`. Then, highlight those three cells, click the square in the bottom-right corner, and drag it down.

What you should observe (after releasing) is that Google Sheets has automatically filled the next handful of cells for you, detecting the pattern you were setting up and filling in the rest. Handy!

This works horizontally as well as vertically - you can auto-fill rows, too, if you just drag horizontally.

1.2.2 Using Cells

As before, though, this really comes into its own when you combine it with the ability to reference cells.

Suppose you have three columns. You want your third column to be the product of your first two - that is to say, `C1` should be `=A1+B1`, `C2` should be `=A2+B2`, etc.

If you enter the first entry in C1 and drag down as before (with even just once cell this time!), it'll automatically increment all cell references. So, in the second row, A1 will turn into A2 and B1 will turn into B2, and so on.

This means that if you're doing a bunch of identical calculations (say, on a bunch of different data points), you can do the calculation once, drag, and get all the results.

If you want a shortcut that's faster than click-and-drag, double-clicking on that same square will automatically fill the column down as far as the previous column goes. For instance, if you had every cell from A1 to A7 filled in and you double-clicked on an entry in B1, it would autofill everything down to B7, just as if you had clicked and dragged to that point.

1.2.3 Fixing a Cell Reference

What if you didn't want them all to change, though? Suppose you wanted your first one to be something like =A1+E7, your second one to be =A2+E7, etc. - you want one of the cells stay fixed when the other one changes.

The way to do this is to use a dollar sign, $. Enter it as =A1+$E$7, and the autofill will not change the E7 anymore.

The first dollar sign fixes the letter, and the second one the number. If you wrote E$7, then the column (E) could change if you autofilled horizontally, but not the row (7) when you autofill vertically. Similarly, $E7 could change 7 but not E.

1.3 Best Fit Line Calculations

In this class, we'll often want to run a line of best fit through some data.

Unfortunately, the default methods that Google Sheets (or Excel) gives for doing this does not give something we want, the uncertainty in the numbers it gives (slope and intercept). So, we'll have to learn how to do it a bit more directly.

To do this, we will use the LINEST function. This function makes a line-of-best fit, but has the ability to tell you more than the automatic fitter does.

Whenever a best fit line is required in this class, there will be a small side-table for it. You will enter the LINEST function into the top-left corner, just below where the units go.

There are four parameters of this function: in the documentation, it is described as `LINEST(known_data_y, known_data_x, calculate_b, verbose)`.

Their meaning is as follows:

- `known_data_y`: This is the data you're using for your y-axis. If this is a range of cells listed vertically (say, `B5` through `B19`), then it will be written as `B5:B19`. This should be filled in automatically if you highlight a range of cells.

- `known_data_x`: Data you're using for your x-axis. Same format as the y-data.

- `calculate_b`: Enter `1` or `TRUE` if you want an intercept for your plot; enter `0` or `FALSE` if you want the best fit line to be forced to pass through the origin.

- `verbose`: This should always be entered with `1` or `TRUE`. This tells Google Sheets that we want all of the information it can give us and not just the slope/intercept, which is how we get the uncertainties.

Thus, if you want to plot `B2` through `B9` as your x-data and `C2` through `C9` as your y-data, with the best-fit line locked through the origin, then you should enter into that upper-left box the command `=LINEST(C2:C9,B2:B9,TRUE,TRUE)`.

This will auto-fill not only the cell you enter it into, but ten cells (two wide, five down) with that as the upper-left cell. There is a variety of data this will output; we will only use the first five cells.

The first two cells (the top row) contain the estimates of your parameters: the first cell is the slope, and the second cell is the intercept. The next two cells (the second row) are the uncertainties in those quantities - the third cell is the slope uncertainty and the fourth cell is the intercept uncertainty.

The fifth cell contains the R^2 value. This *correlation coefficient* is a measure of how much your data "looks like a line": if it is close to 1, your data is highly linear.

The remaining cells contain various statistical information that we will not be using. If you are curious, though, you can find out what they are telling you at `https://support.google.com/docs/answer/3094249?hl=en`.

2 Making Plots

2.1 Why We're Not Using Google Sheets For Plots

Given that we're doing all of our data analysis in Google Sheets, it might seem silly that we're not using it to make our plots too.

Unfortunately, as of this writing, Google Sheets does not have certain functionalities that we need for this class. In particular, one cannot make customizable x and y error bars (that is to say, error bars that vary from data point to data point).[4]

While Excel and other spreadsheet software *can* make such plots, we are trying to avoid requiring any software you would need to download. Therefore, we will use a standard plotting tool made specifically for this course.

If you want to use Excel or another spreadsheet software instead, talk to your TA. They may allow it, but if you choose to do this, you should be sure your plot looks correct (compare to the kind of plot the plotting tool makes).

2.2 Using the Plotting Tool

The easiest way to use the plotting tool is to use the manual interface available at `http://phylabs1.physics.sunysb.edu/~physlab/Summer2018/PlottingTool/PlottingToolInterface.html#manualinterface`.

This interface should be straightforward to use - enter your data and other information, and click "Make the plot!"

Unfortunately, there are a few limitations of this version. First, you have to manually enter all of your data - rather annoying, even if you copy and paste each cell. Second, there's no way to save the data you have - if you close the page, you'll need to re-enter all of your data.

In the interest of solving these problems, we added another way to use the plotting tool, which avoids those issues:

1. Make a copy of the "Plotting Tool Data Table," a Google Sheets template available at `https://docs.google.com/spreadsheets/d/1soepnWCl1uvqjlCq0j7T9RFqNAJqHlY4UYQ-I`
2. Copy your data into that sheet. Since, like your data table, it's a Google Sheet, this should be mostly be a matter of copying-and-pasting.[5] This will be your saved, easily-editable backup copy.
3. Download this data table as a `.csv` file. Click on `File ▸ Download as ▸ Comma-separated values (.csv, current sheet)`.
4. Go to `http://phylabs1.physics.sunysb.edu/~physlab/Summer2018/PlottingTool/PlottingToolInterface.html#automaticinterface` and upload that `.csv` file, click "Make the plot!"

[4]Technically, there are ways to make it work, but it takes an absolutely ridiculous amount of finagling, requiring making every data point its own series, and so we won't do that.

[5]Don't worry if this messes up the background colors - those won't matter.

While slightly more technically complicated, it can make your life easier if you learn how to do this, since you won't have to re-enter your data every time.

The plotting tool will also provide a fit. This should agree with the results of the `LINREG` procedure described in section 1.3; if it does not, check that you copied all of your data correctly, and chose to set the fit through the origin (or not) in the same way.

3 Specifics to This Lab

3.1 Getting the Data Sheets

You can find a folder with read-only versions of all of the data sheets at `https://drive.google.com/drive/folders/12bNUcIVG4GQTPsAlFcWwOVXsauVaAbE9%22`.

In order to edit them, you will need to save a copy. This appears in the drop-down menus as `File▸Make a Copy`. After doing this, you should be able to edit the cells and enter in numbers of your own.

If you ever find that you messed up the formatting my accident, you can go to the original read-only sheet, copy everything (Ctrl+A, Ctrl+C), then go back to your copy, go to the first cell (A1) and paste only the formatting using Ctrl+Alt+V, or by right-clicking and using `Paste special▸Paste format only`.

This will copy all the background colors, cell borders, etc. - everything except the contents of the cells. This is handy because it won't overwrite anything you have written in your sheet, only the formatting.

3.2 Submitting the Data Sheets

Click `Share` in the upper-right corner, and a menu will pop up.

Click `Get sharable link` in the upper-right corner of that menu.

Copy the link and paste it into the submission box in Blackboard, and your TA will now be able to view your data sheet.

You should *also* submit a copy of your data sheet with the rest of your report (hard-copy or PDF, depending on whether you turn it in physically or electronically).

Your TA may request that you change the "Anyone with the link **can view**" to "Anyone with the link **can comment**." (This will enable your TA to tell you what you did wrong, if anything, directly on your Google Sheets document.)

3.3 Colors on the Data Sheet

The data sheets are color-coded for your convenience, so you can easily identify what is necessary to do in-lab and what you can finish later. The colors are as follows:

- Yellow: Necessary to do in lab.
- Blue: Calculation you can do outside of lab time, if necessary. Not final results, but helpful things to have along the way.
- Purple: Key final result of your calculation.
- Red: You can ignore these cells. (Typically covers neglected uncertainties or unused parts of the best fit line analysis.)
- Pink: Options (for the plotting tool).

3.4 Dropdowns

Some data sheets will have drop-down menus in some cells. Click the little arrow to select an option.

You can also enter in text yourself if (for some reason) none of the options match your observation/answer.