

Determination of Density: Analysis

B. D. Lamp, D. L. McCurdy, V. M. Pultz and J. M. McCormick
Copyright ©2020 Truman State University

Introduction

Not so long ago a statistical data analysis of any data set larger than a few points was a time-consuming and tedious procedure.^{1,2} This was changed first by the introduction of personal computers and then by spreadsheets, which are computer programs that allow the user to enter and manipulate numerical data. Spreadsheets were originally designed for business applications, but have become essential tools for data analysis in all of the sciences because of the ease with which they can perform complex calculations and graph the results. Many hand-held calculators can perform similar tasks, but spreadsheets have the advantage because they store data in an easily edited form and produce higher-quality graphs.


In this exercise you will learn the basics of statistical data analysis and of spreadsheet operations using the program Excel. The data that you will manipulate will be measured values of copper's density obtained by first measuring a copper block's dimensions and then by water displacement. While performing this exercise, refer to [Introduction to Statistics in Chemistry, Preparing Graphs](#), http://vpultz.sites.truman.edu/files/2018/08/Excel_2013_exercise_2018Aug.pdf, and [Guide to Excel2010](#). Save your work frequently.

IMPORTANT! Either have Microsoft Excel installed on your computer or access Excel on a university computer or through <https://view.truman.edu/>. Screen shots in this experiment are from Excel 2010, but other versions of Excel are similar. Other spreadsheet programs are acceptable if they have the same capabilities for formatting, calculations, and graphing. The “help” function in Excel or other online resources are useful if you run into problems locating specific features.

Results and Analysis

Graphical Determination of Density using a Ruler to Measure the Volume

Open the class data for your lab day in Excel and set up the first worksheet as shown in Fig. 3. In the cell where it says *Names*, enter your name. The class data for the copper blocks is in columns A through E. Only numbers are in these cells; units are in column headings. In cell F3 you may enter “Volume (cm³)”, in cell G3 “DV (cm³)” and in cell H3 “DV/V”. Use *F_ormat*, *C_ells*, as required, to subscript or superscript text, or to change text to symbol font. Alternatively, select what you wish to make as a superscript, make sure you are on the Home tab, click on the lower right corner of the Font panel, and check Superscript. Δ could also be inserted by clicking the Insert tab, choose Symbol at the far right, go to Greek and Coptic subset, and click on Δ.

All significant figures must be shown! Spreadsheets drop trailing zeros, even if they are significant. Since all of the length, width, and height measurements were measured to the nearest 0.01 cm, it is quickest to select the block of these measurements (cells B4 through D24 or D22, depending on how many blocks were used) and click on the increase decimal places button ( in the Number group) to display two digits after the decimal in all of these cells. Adjust all mass measurements to three digits after the decimal.

Enter a formula in cell F4 to calculate the volume of the block from its dimensions. The volume of a rectangular box is its *length* x *width* x *height*. The first block's length, width and height are in cells B4, C4 and D4, respectively. To calculate the volume in cell F4 enter “= B4*C4*D4” (without the quotes) in cell F4.

Copy this cell and paste it into cells F5 through F22 or F24, depending on how many rows have data.

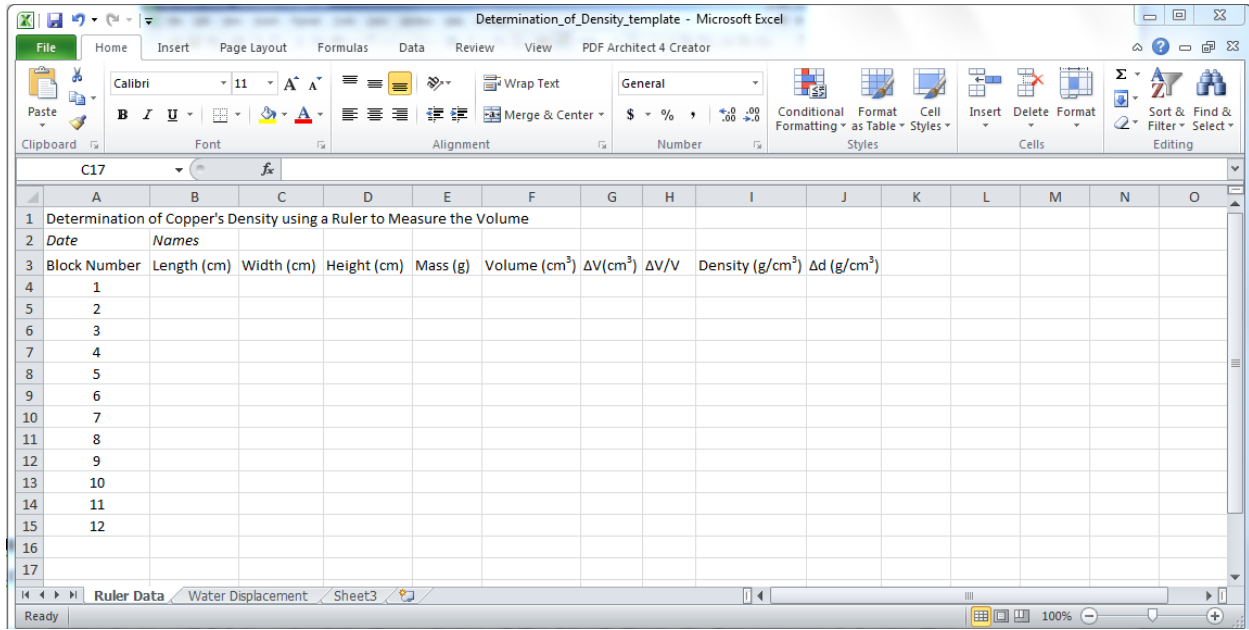


Figure 3. Worksheet for the data from the determination of copper's density from the block's dimensions and its mass.

Determine the uncertainty in the volume, ΔV , from the uncertainties in each dimension (Δx , Δy , Δz) using Equation 1 (Eqn. 1). Since the dimension measurements were estimated to the nearest 0.01 cm, Δx , Δy , and Δz are all 0.01 cm. Write the equation and values used in your lab notebook.

$$\Delta V = \pm V \sqrt{\left(\frac{\Delta x}{x}\right)^2 + \left(\frac{\Delta y}{y}\right)^2 + \left(\frac{\Delta z}{z}\right)^2} \quad (1)$$

Translate Eqn. 1 into Excel format and enter it in cell G4. Multiplication is '*', division is '/', powers are '^' and the square root function is SQRT. So, you will enter in cell G4 “=F4*SQRT(((0.01/B4)^2+(0.01/C4)^2+(0.01/D4)^2)” (without the quotes).

Copy and paste cell G4 into cells G5 through subsequent cells which will refer to data.

The uncertainty in the volume has one significant figure since even that one digit is, by definition, uncertain. Select volume and ΔV pairs and format to display appropriate significant figures. How does this approach compare with the simple rules for significant figures?

In cell H4 enter a formula to calculate $\Delta V/V$ (the uncertainty in the volume measurement divided by the volume). Copy and paste cell H4 into cells H5 through H22 or H24. Set the number of significant figures in column H to two by adjusting the number of digits after the decimal. (Two is not the correct number of significant figures, but it will be useful later on.)

In cell I4 enter a formula to calculate the density, d , of the first block. In cell J4 calculate the uncertainty associated with this single measurement of the density, Δd , using Eqn. 2 and cell references to values already calculated and assuming that Δm is ± 0.001 g. Select both cells I4 and J4, then copy and paste into subsequent rows. Select density and Δd pairs and format to display appropriate significant figures. Record the calculations for your block in your laboratory notebook.

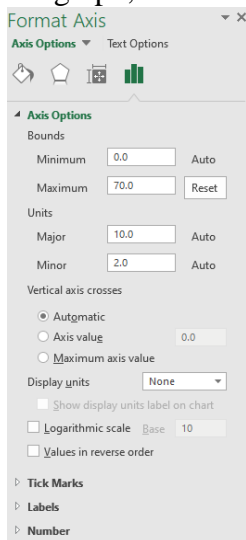
$$\Delta d = \pm d \sqrt{\left(\frac{\Delta V}{V}\right)^2 + \left(\frac{\Delta m}{m}\right)^2} \quad (2)$$

Prepare a graph of the copper blocks' volume as a function of their mass in Excel. We need to find the relationship between the block's mass (m) and its volume (V). We assume that there is a linear relationship, which we can write as Eqn. 3, where a is the slope of the line and b is the y-intercept. To avoid confusion, the slope is given the symbol " a ", instead of the usual " m ",

$$V = a \cdot m + b \quad (3)$$

because the mass already has that symbol. Prepare a graph with volume as the dependent variable and mass as the independent variable. Since density equals mass divided by volume it would seem logical to graph mass as a function of volume so that the resulting straight line's slope would be the density. Although this looks good, it is wrong. This is because when one performs a linear regression on a data set the quantity with the smallest error must be on the x -axis. If this is not done then the results may be statistically meaningless. In our case we know the mass much more precisely than we know the volume and so the mass must be on the x -axis.

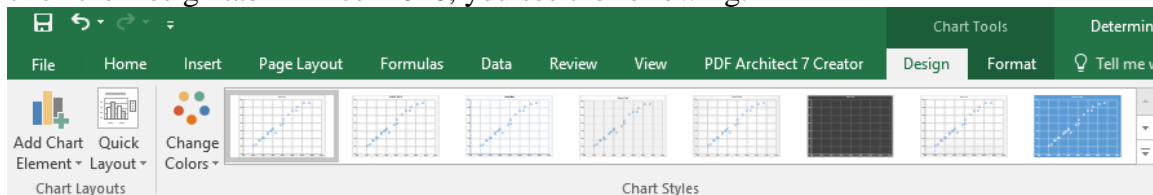
As soon as you make the graph, it appears on the sheet with the source data. Immediately click on Move Chart at the upper right, click New Sheet, and click OK. Right-click on a value on the x -axis, Format Axis, change the maximum to 70 from 80 to help the data points fill up more of the graph, click on the triangle to the left of Number at the bottom of the menu at the right,



and change the Decimal places to 0. Hit Enter and close the menu.

Also change the Decimal places to 0 on the y-axis. Remove the gridlines. One way is to right-click on a gridline and choose Delete.

Insert a trend line in the graph. Left-clicking on the area with points makes it active. Click the plus sign at the upper-right corner. Move the cursor to Trendline and click on the triangle to the right and choose More Options. Keep the default of Linear. Check the bottom two boxes for Display Equation on chart and Display R-squared value. (Review Graphing Exercise #1 at http://vpultz.sites.truman.edu/files/2018/08/Excel_2013_exercise_2018Aug.pdf if you need help.) If you click the Design tab in Excel 2016, you see the following:



Add Chart Element is a way to access Axis Titles and other features. The x-axis title should be Mass (g) or Mass, g. The y-axis title should be Volume (cm³) or Volume, cm³. Use the same style for both axis labels. Enter a title that is not redundant. The title may not use the words mass of volume. Choose a title that indicates the purpose of the graph.

Click on the Insert tab and choose Header & Footer. Click Custom Header. In the Left section enter CHEM 130, in the Center section enter your name, and in the Right section click the Insert Date button.

Print one copy of the graph from Excel. It will be a full page. Go to the Ruler Data sheet. Click the Page Layout tab; under Gridlines, make sure boxes for View and Print are both checked. Insert a Header as you did for the graph. Click Orientation and choose Landscape. Print one copy of the ruler data. The student names do not need to be included.

Using Excel's regression package (see page 8 at [Guide to Excel 2010](#)), calculate the slope, the intercept and the uncertainty in the slope and intercept at the 95% confidence limit for these data. Be sure that the slope and intercept found using the regression package are the same as those determined from the trend line. If they are not exactly the same, there is a problem somewhere that you must correct before continuing. Adjust the width of the columns on the regression-output worksheets so that all of the headings can be read. (Try *AutoFit Column Width* under *Format* in Cells group on the Home tab.)

Determine copper's average density, d , from the slope of the best-fit line on your Excel graph. (Do the calculation in the regression output sheet using a cell reference to the slope.) Calculate the uncertainty in copper's density at the 95% confidence limit. First find Δd from the standard deviation in the slope of your Excel graph (Δa , labeled "Standard Error" in the output for the regression package) using Eqn. 4. This must be done because the slope is $1/d$, and so Δa is not Δd .

$$\Delta d = \pm d \sqrt{\left(\frac{\Delta a}{a}\right)^2} \quad (4)$$

When dealing with two-dimensional data sets, we must use a slightly different approach to calculate the confidence interval, as compared to one-dimensional data. In the two-dimensional case, the confidence limit is calculated as $\Delta = t \cdot \sigma$, where σ is the standard deviation and t is determined based on $n - 2$ degrees of freedom, where n is the number of x, y pairs in the dataset. Since we want uncertainty in the density, σ corresponds to Δd . (Notice the absence of the square root of n term, this is not a typo!) Use the TINV function in Excel to find the appropriate value for Student's t . Perform all calculations in Excel and record in your lab notebook. Calculate a percent error for the average value using the accepted density of copper (8.96 g/cm^3).³

On the Page Layout tab, set the Orientation to Landscape and make sure that the table fits onto one sheet of paper by using Print and viewing the preview. Never split tables over two sheets unless absolutely unavoidable. Print one copy. Save the Excel file and record the name you saved it under in your laboratory notebook.

Determination of Density by Water Displacement

Prepare the second worksheet in your Excel file for the water displacement data so that it looks like that shown in Fig. 4. (Your professor went to the worksheet renamed "Ruler Data" and selected all block numbers in column A, then copied and pasted into the worksheet named "Water Displacement" starting with cell A5. He also copied and pasted the masses from the worksheet with the ruler data.) The reported values for the volume of water displaced will generally be different from the volumes calculated based on ruler measurements. Check the format of the masses to verify that they still represent measurements made on a balance which reads to the nearest 0.001 g. Since the volumes of water displaced will be read to the nearest 0.1 mL, format the volumes to reflect this precision.

Calculate the density of each block from its mass and volume by entering the correct formula in cell D5 and copying down the column. In cell G5 enter a formula to calculate the uncertainty in the density for this particular block using Eqn. 2 and remember that ΔV will be 0.1 mL since you used a graduated cylinder marked off every 1 mL. Copy and paste for the rest of the blocks. Note that the densities and their uncertainties for individual blocks usually differ from what you calculated based on ruler measurements. Since the uncertainty has one significant figure, format the density and uncertainty in the density so that each pair for a particular block has the same number of digits after the decimal.

	A	B	C	D	E	F	G
1	Determination of Copper's Density using Water Displacement						
2	Date	Names					
3			Volume Water				
4	Block Number	Mass (g)	Displaced (mL)	Density (g/cm ³)	$x_i - x_{avg}$ (g/cm ³)	$(x_i - x_{avg})^2$ (g ² /cm ⁶)	Δd (g/cm ³)
5	1						
6	2						
7	3						
8	4						
9	5						
10	6						
11	7						
12	8						
13	9						
14	10						
15	11						
16	12						
17			$x_{avg} =$		g/cm ³		
18			$\Sigma(x_i - x_{avg})^2 =$				
19			S =				
20			STDEV =				
21			$\Delta =$				
22							

Figure 4. Worksheet for the data from the determination of copper's density by water displacement.

Inspect the data to identify whether any point seems to be out of place. If you find a point that you think is an outlier, first check that there were no computational or other gross errors, then perform the Q -test ([Introduction to Statistics in Chemistry](#), Eqn. 6 and Table 2) on the suspect point. The easiest way to perform the Q -test is to select cells A5:G16 in the example above but A5:G23 for the Th lab and A5:G25 for the Tu lab, click on the *Data* tab, then *Sort*, unclick box My data has headers, and choose *Sort by Column D*. Calculate Q for the lowest and highest densities in column H. Actually, we have more data points than one individual will often have. For large data sets ($N > 10$) a data point that lies more than 2.6 times S (or σ) from the average may be excluded. Only one data point may be excluded; remove the worst. If you can eliminate one point, format it in Strikethrough and indicate in worksheet that this point was “eliminated on the basis of a Q -test.”

In the first empty cell below the last calculated density, cell D17 in Figure 4, calculate the average density using the AVERAGE function. Adjust the number of significant figures in the calculated average (remember that the average can be no more precise than the least precise number used to calculate it). If you eliminated a data point, do not include it in the average or any other calculations.

Calculate the estimated standard deviation, S , of the data using Eqn. 5, where N is the number of points in the data set, x_i is each individual measurement and x_{avg} is the average. First, we will do it in a step-wise calculation, and then use a built-in Excel function.

$$S = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - x_{avg})^2} \quad (5)$$

In cell E5 enter a formula to calculate $(x_i - x_{avg})$ using an absolute reference to the cell which contains the average. An absolute reference is denoted using the '\$' character. In Figure 4 you would enter (without the quotes) “=D5-\$D\$17” (where \$D\$17 is the cell containing the average, so \$D\$22 for Th lab and \$D\$24 for Tu lab) to calculate the difference between each data point and the average. This will stop Excel from automatically changing the reference to the cell with the average when we copy and paste cell E5 into subsequent cells in column E. Look at the equations in the cells after you paste them in to convince yourself that what you expect to happen is actually happening. Adjust the significant figures as needed. (All will have the same number of digits after the decimal.)

Enter a formula in cell F5 to calculate $(x_i - x_{avg})^2$, and then copy and paste it into the appropriate cells in column F. To raise a number to a power in Excel use the '^' character. So, in cell F5 enter (without the quotes) “=E5^2” to square the contents of cell E5. Be sure that each cell displays close to the correct number of significant figures.

Figure 4 needs a formula in cell D18 to calculate $\Sigma(x_i - x_{avg})^2$ using the SUM function. In the cell to the right of the formula enter the units for the numerical value. If you removed a data point based on the *Q*-test, do not include it in the range for the SUM function.

Figure 4 needs a formula in cell D19 to calculate *S* from cell D18 using the SQRT function to implement Eqn. 5. Treat the square root as you would treat division to determine the number of significant figures. The number of data points, *N*, is an exact number, and as such does not affect the number of significant figures in the standard deviation. Note that, as with the average, the standard deviation cannot be more precise than the original data! Figure 4 has you enter the units of *S* in cell E19. Adjust for your data set.

According to Figure 4, in cell D20 use the STDEV function to calculate *S* directly from the data in cells D5 through D16. Enter the units for cell D20 in cell E20. Adjust these cell references to what is appropriate for your data set. Note that from now on, if you need to calculate any statistical quantity (i.e., average or standard deviation) you can use the spreadsheet's built-in functions, instead of the elaborate procedure that you used in this lab.

According to Figure 4, enter an equation in cell D21 to calculate the uncertainty at the 95% confidence level from the estimated standard deviation in cell D21. The expression that you need is Eqn. 5 at [Introduction to Statistics in Chemistry](#). Use the TINV function to find the appropriate value of Student's *t*. In cell E21, enter the units of Δ .

Print one copy of this worksheet (Landscape orientation with gridlines printed). Calculate the percent error in copper's density as measured by this method. Save the Excel file again.

Conclusions

Include a discussion of whether the assumptions made in each method were likely to be valid, and how the results would be different if the assumptions were not valid. This discussion can be qualitative (i.e., no calculations required), but it must address all of the assumptions.

When you determined the density by measuring the copper block's volume using the ruler you calculated $\Delta V/V$, the ratio of the error in the volume to the total volume. Include in your discussion what you notice about $\Delta V/V$ for the small blocks as compared to the large blocks. What does this mean and why might this be a problem? How would you modify the experiment to minimize its effect on the density?

References and Notes

1. Schlotter, N. E. *J. Chem. Educ.* **2013**, *90*, 51-55. This article is available as a PDF file at <http://pubs.acs.org/doi/pdf/10.1021/ed300334e> for Truman addresses and *J. Chem. Educ.* subscribers.
2. Jolly, W. L. *Encounters in Experimental Chemistry*, 2nd Ed.; Harcourt Brace Jovanovich: New York, 1985, 52-54.
3. *CRC Handbook of Chemistry and Physics*, 64th Edition; Weast, R. C., Ed.; CRC Press: Boca Raton, FL, 1984, p. B-11.