# Lab Report Instructions for Experiment #2

Setup

1. Log into the RStudio Server (rstudio.reed.edu) with your Reed ID and password
2. Welcome to RStudio. I will give you a brief introduction to where things are on your screen before beginning your first report.
3. Create a New Project
    a. File -> New Project _> New Directory - > Empty Project
        i. Leave the git repository and packrat options unchecked.
    b. Directory Name: CHEM101_E2_Names
4. Make sure you are in your Project. It should be visible in the circled section below.
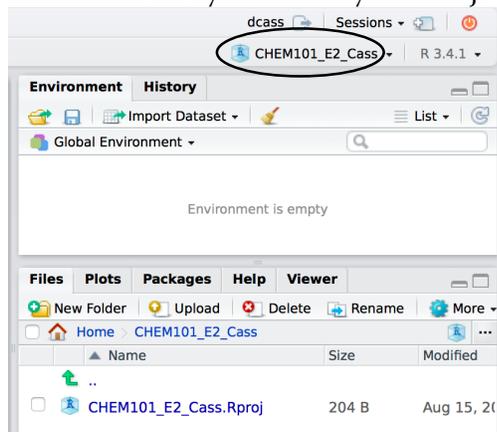
5. Share your project with your partner(s) so that everyone can work on and modify the document.
    a. File -> Share Project
    b. Type in the Reed user name of your partner(s) in the dropdown box (For example my user name is dcass)
    c. Click Add
    d. After you have shared the project with everyone, click OK.
    Note: There is also Project URL below the add list. Sharing this URL with each other results in a quick way for people to get to the file that has been shared with them.
6. Open the E2 Beers Law Template
    Note: Templates are R Markdown files that give you an outline for your lab reports. Sometimes I have included important code for you. Sometimes the code is available in the instructions or from a previous lab and therefore I have simply left space for it. Anything that has a # before it is commented information for you. Please read these notes.
    a. Go to: File -> New File -> R Markdown
    b. Choose "From Template", highlight E2 Beers Law, click OK
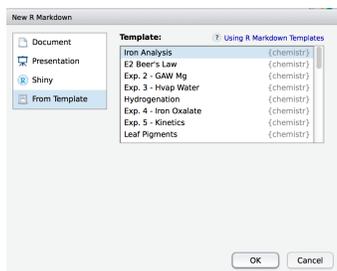
7. The template opens as an unsaved document which will become your lab report. The first thing you will want to do is save the document.
   a. Go to File -> Save As
   b. Name your document (ex. E2_PartnerA_PartnerB) and make sure you save it in your CHEM101_E2_Names Folder.

An overview of the R Markdown Template

1. What is an R Markdown File?
   a. R markdown is a document that allows you to combine R code (aka data analysis) within the actual written document.
   b. A R markdown file allows reproducible data analysis.
   c. A R markdown file has the extension .rmd
   d. There are three different parts of a .rmd file
      i. Markdown = the text portions of the file
      ii. R Chunks = the coding, data analysis portions of the file
      iii. YAML header = controls how your document appears in its final format (in our case as a Word document)
2. YAML Header
   a. Lines 1-6 below (an on your template) are the YAML header. The only things you will need to do for this section is to change the title, name, and date. Make sure when you replace these items, you keep the quotes around the title, name, and date.
   b. The output is already configured based upon what we would like for CHEM101.

3. R Chunks
   a. R Studio will compute and analyze any code that is embedded in an R Chunk. These chunks can be seen because they are tan and start with ```{r} and end with ```. See line 8 and 10 in the figure above.

b. The first R chunk in our templates will always be what is seen in lines 8-10 above. R is an open source program and therefore a lot of people have created add-ons to make R more useful. We have combined what others have done and things we want into a package called chemistr. This first chunk tells R Studio to use this package.

c. There are also sometimes addenda after the R. For example:
   i. In line 8 you see include = FALSE, this means the chunk should be evaluate but neither the code nor its output is displayed.
   ii. In line 16 you see echo = FALSE indicating the code will not be shown but the results or output will be displayed.
   iii. In line 26 you see fig.cap = "...", this allows you to add a figure caption to a graphical result.
   iv. There are many other options, but these are most important for today.

d. I have also added notes or comments within the R Code which are not analyzed. This is done by using the # sign before the note and often contains explanatory information, such as found on line 17.

e. R Chunks can be run as you create them by clicking the green arrow at the top of a chunk. See the circled arrow below. Try running the first chunk.
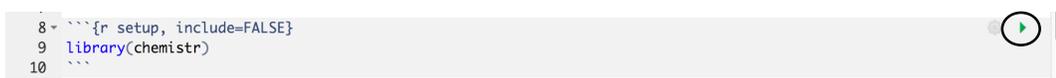
```
8 ```{r setup, include=FALSE}
9 library(chemistr)
10 ```
```

f. You can now see in the History tab on the right side that the command has run. You can also see that in your packages tab below, the chemistr package is checked. Finally, in the Console window, you can see that packages bundled with chemistr were also added.
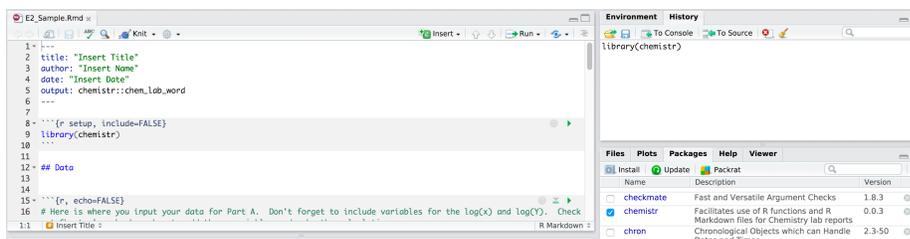
4. Markdown
   a. The white space in your .rmd file is space in which you can write plain text. This is done using the Markdown annotations. Markdown allows you to format in many different ways. Help for this can be found in a variety of locations on-line. A quick resource is the Rmarkdown Cheatsheet found at https://www.rstudio.com/wp-content/uploads/2016/03/rmarkdown-cheatsheet-2.0.pdf
   b. Note: You will be rendering the file into a word document, so if you are having trouble formatting in Markdown you may do the formatting in word instead.

c. You can see that the first text line is on line 13. The ## before Data classifies it as a header, level 2. The output in the word file will look like:
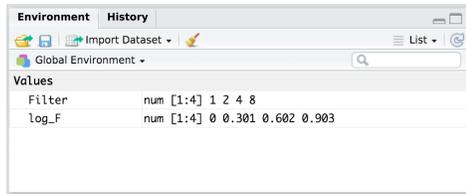
Data

Experiment 2 Lab Report

1. The purpose of this lab is to determine what mathematical relationship exists between the transmittance of light and two variables (number of films and concentration of molecules in a solution). This is done by graphical analysis. So we are going to have to learn how to do the following things in R:
   a. Input data (called vectors in R)
   b. Create a scatter plot
   c. Fit the scatter plot data to a straight line
   d. Extract the fit data for analysis
2. Inputting Data into R
   a. R data will always be entered into a chunk. There are a variety of ways the data can be added, but for our data we will always be manually inputting the values. If the data were already in an electronic format one could upload the database directly.
   b. In order to use this data for later manipulations, the data must have a name. To name a vector, you use the format:
      Name <- c(Data1, Data2, Data3, etc)
   c. Notes on naming vectors:
      i. Capitalization counts, so if you name something Var then you must use the capital V whenever referring to that variable.
      ii. Don't start with numbers. For example: 2D will not work as a name, but D2 is fine.
      iii. Don't include spaces, use an underscore. For example: data 2 will not work as a name, but data_2 is fine.
   d. Input of Part A – Transmittance versus Film Number Data
      i. Lines 18 and 20 have vectors already named for your raw data from part A. Replace the values for these vectors with your data.
3. Mathematically Manipulating Data in R
   a. The R chunk and the R Console can be used as a calculator, similar to how you can do math in Excel to an entire column of values.
   b. In your console window find the log(100) by typing in log10(100). Note: If you type log(100) it will take the natural logarithm of 100 (ln(100)).
   c. Now you can do that same math on all of the film values by typing, on line 19,:
      log_F <- log10(Film)
   d. Make sure your code is working by highlighting lines 18 and 19. Then click
      ⟶ Run ▾. Select "Run Selected Lines". *Note that you can choose how much of your code to run at a time based upon what you highlight or choose.* You will see these new vectors appear in the Environment tab.

e. You can also see the results of these vectors by typing their names in the console.

f. Now create and define the new vector log_T1, which is the log10 of your transmittance values, on line 21 in a similar way to what was done in line 19.

4. Data Frames

a. The last step in creating our data is making a data frame. A data frame is basically a data table or 2-dimensional array of variables, that are **all the same length**. If you try to combine vectors of different lengths, the data frame code will give you an error.

b. Line 22 creates a data frame using the code data_frame followed by the vector names that you want to combine.

c. Since this is the last step in this chunk, click the green arrow to run the entire chunk.

d. You should now see all of your variables in the environment window and also the data frame film_data. You can look at the data frame two ways:

    i. Click on the miniature spreadsheet next to the film_data in the environment window. This will create a new tab showing the array.

    ii. In the Console, type in film_data and press enter. The array will appear in the console.

5. Creating a Graph

a. Now that we have all of our Part A data, we need to create a scatter plot of our data. The first model we will look at is the linear relationship between Transmittance and Number of Film Slices. Look back at the mathematical models information from your lab instructions. If there is a linear relationship, then a good fit to the data is Transmittance = Slope*(Number of Film Slices) + Intercept. This means that we want to plot Transmittance on the Y axis and Film Slice on the X axis.

b. The scatter plot is created on line 28, using the code chem_scatter( ). The information that is needed to complete the graph is:

    i. data = name of the data frame containing the variables to be plotted

    ii. xvar = name of x-variable

    iii. yvar = name of y-variable

    iv. xlab = label wanted on the x axis (including units)

    v. ylab = label wanted on the y axis (including units)

    vi. This code also automatically best fits the slope and intercept. It also includes the best-fit line on the graph.

c. Highlight line 28 and run this line by clicking [Run]. You will see your graph appear below the R chunk.

6. Fitting and Extracting Data from a Plot
    a. Now that you can see the best fit line on the graph, we want to extract the parameters from the straight line fit of this linear model. The code for this is lm( ). The information needed to complete this code is:
        i. Linear Formula = The format for this is y-variable ~ x-variable where the fit will include variables for slope and intercept.
        ii. data = Name of the data frame that contains the variables
    b. To see the fit parameters of fit1 the code summary(fit1) is included. Putting this at the end of the code will include these results in the final document. If you didn't want the entire summary in the document, you could simply run the code and then type summary(fit1) in the console.
    c. Run this entire plot 1 chunk to see your graph and fit parameters. You should now see both the graph and an R Console window below this chunk. Record the slope (#1), intercept (#2), and R^2 value (#3) from the summary in your lab notebook. You will need these values later.
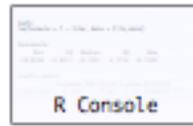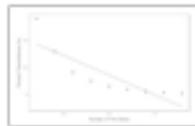

```
R Console

Call:
lm(formula = T ~ Film, data = Film_data)

Residuals:
     Min      1Q   Median      3Q      Max
-10.0386  -8.0911  -0.5161   4.3714  18.5989

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) #2 42.979      7.447    5.771 0.000683 ***
Film        #1 -5.768      1.323   -4.358 0.003322 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 10.25 on 7 degrees of freedom
Multiple R-squared:  0.7307, #3 Adjusted R-squared:  0.6922
F-statistic: 18.99 on 1 and 7 DF,  p-value: 0.003322
```

7. Finishing the data analysis for Part A.
    a. Now that you have analyzed the linear relationship, refer back to the E2 Lab Instructions and analyze the exponential and power model for transmittance versus number of films. Use similar code to the above example, just make sure you use different variable names so you don't overwrite the existing values. You will need to graph these models in a linear manner and extract the fit parameters from each graph.
    NOTE: There is a mistake in the info on line #35 and #36, for a logarithmic model you want to plot log(Y) vs x NOT log(x) vs Y.

      b.   Once you have the fit parameters for all three models, decide which model fits the data best. Write an equation for this model in your lab notebook. Have your TA or instructor take a look at your model before moving onto part B.

8.   Finish the data analysis for Part B.
      a.   Using what you did for Part A as a template, input the code for Part B and the analysis of the best model to describe the relationship between transmittance and concentration.
      b.   Once you have finished running all the code and analyzing the results, write the equation for the model which best describe the transmittance versus concentration relationship in your lab notebook. Have a TA or Instructor look at this before you continue.

9.   Constructing a Table
      a.   In your notebook you have R^2 values for several different models. It will be useful in your discussion to refer back to those numbers and therefore placing them all in a table will be helpful.
      b.   First of all you need to enter all the data into the vectors and create a data frame. The names of these vectors and the data frame have already been provided.
           i.   Note that the vector Model does not contain numerical data, but instead contains text. When the vector contains text, each value needs to be surrounded by quotes.
          ii.   After imputing all the data run this chunk. You should see your new vectors and data frame appear in the environment window.

      c.   In the next chunk the table is created using the code chem_table( ). The information needed to complete the code is:
           i.   data = data frame you want to make into a pretty table
          ii.   caption = The text you want to display above the table as a caption. This should be in quotes. For example: "Table 1: A Lovely Caption".

10. Discussion
      a.   The discussion can be written in R or in Word. It is useful to write it in R using the Markdown formatting language and then use Word to do the final formatting and cleanup. This text should follow the format as described in your lab instructions.

11. Final step in R
      a.   Your final step in R is to export the Word document. To do this you will click the knit button ![Knit] . Then it will ask you to either click download file or ok. Choose download file. Then open this new file in Word.

12. Final Edits and Submission
   a. Proofread your document, making any changes that are necessary in Word. This will include spellcheck, formatting, and significant figures.
   b. Once everyone in your group is happy with the lab report then you should save it as a PDF. The PDF should then be opened and made sure it looks correct.
   c. Your last step is to have one person in your group upload the PDF to the Lab 2 – Beer's Law Lab Report assignment on Moodle.