

CHESTER ISMAY AND DANIELLE CASS

LEARNING R IN CHEMISTRY AT REED COLLEGE

Contents

<i>Introduction</i>	5
<i>What is this?</i>	5
<i>Starting the labs</i>	6
<i>Requesting help</i>	8
<i>1 Lab 1: Light</i>	11
1.1 <i>Getting the lab template</i>	11
1.2 <i>The parts of the lab template</i>	11
1.3 <i>Uploading and inserting external pictures</i>	14
1.4 <i>Note on white space</i>	14
1.5 <i>Spell-check</i>	15
1.6 <i>Converting your Word document to PDF</i>	15
<i>2 Lab 2: Beer's Law</i>	17
2.1 <i>Getting the R Markdown lab template</i>	17
2.2 <i>The parts of the lab template</i>	17
2.3 <i>Note on white space</i>	19
2.4 <i>Spell-check</i>	19
2.5 <i>Converting your Word document to PDF</i>	20
2.6 <i>Note on requesting help</i>	20
<i>3 Lab 3: Reactor</i>	21
3.1 <i>Getting the R Markdown lab template</i>	21
3.2 <i>The parts of the lab template</i>	21
3.3 <i>Note on white space</i>	23
3.4 <i>Spell-check</i>	23

3.5	<i>Converting your Word document to PDF</i>	23
3.6	<i>Note on requesting help</i>	23
4	<i>Lab 4: Iron</i>	25
4.1	<i>Getting the R Markdown lab template</i>	25
4.2	<i>The parts of the lab template</i>	25
4.3	<i>Note on white space</i>	26
4.4	<i>Spell-check</i>	27
4.5	<i>Converting your Word document to PDF</i>	27
4.6	<i>Note on requesting help</i>	27
A	<i>RStudio Desktop</i>	29
A.1	<i>Downloading R and RStudio Desktop</i>	29
A.2	<i>Downloading and installing chemistr</i>	29
B	<i>Lab 1: Light (More Details)</i>	31
B.1	<i>The <code>chem_table</code> function</i>	31
B.2	<i>The <code>include_graphics</code> function</i>	32
	<i>References</i>	32

Introduction

What is this?

In the HTML version of this book, you can also download the PDF version of the book by clicking on PDF button in the top toolbar of the page. HTML is the preferred format but the PDF format may be preferred for some readers. Links to the different GIFs directly found in the HTML version are provided in the PDF version.

This resource is designed to assist students in CHEM 101/102 in using RStudio and R Markdown to complete their labs. (A more general reference to the specifics of R, RStudio and R Markdown is available in a different free book here (Ismay, 2016)). This resource will show you GIFs explaining how to do some common procedures you'll need to do to complete labs. In addition, it will provide more details about the specifics of the different lab templates available in the **chemistr** R package. Each chapter of this book will correspond to each of the different labs.

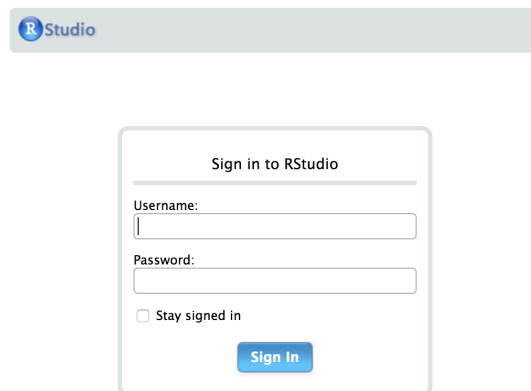
Lastly, if you are interested, you'll find instructions on how to complete the labs using RStudio Desktop instead of the RStudio Server in the Appendix. The Appendix will also include descriptions of the code that is sitting "behind the scenes" in many of the functions created in the **chemistr** package: **chem_table**, **chem_scatter**, **chem_clustered.bar**, etc. These functions were written to ease a beginner into working with R. Those interested in customizing their lab reports further should look over the R code in the Appendix. Additionally, many of the needed packages are automatically loaded when the **chemistr** package is loaded using **library(chemistr)**. You'll see how each of these packages is used in the Appendix.

If further clarification is needed on any other aspect of the book, please email me with a reference to the error/area where more guidance is necessary. More advanced users are encouraged to create a GitHub issue here. Pull requests on GitHub for typos or improvements are also welcome and you can easily do so by clicking on the Edit button near Search at the top of the HTML version of the book.

This book will evolve and be updated as needed based on feedback. You can always check the date at the end of the chapter to see when the book was last updated. It is recommended that you use Google Chrome as your browser since GIFs sometimes do not load automatically with other browsers.

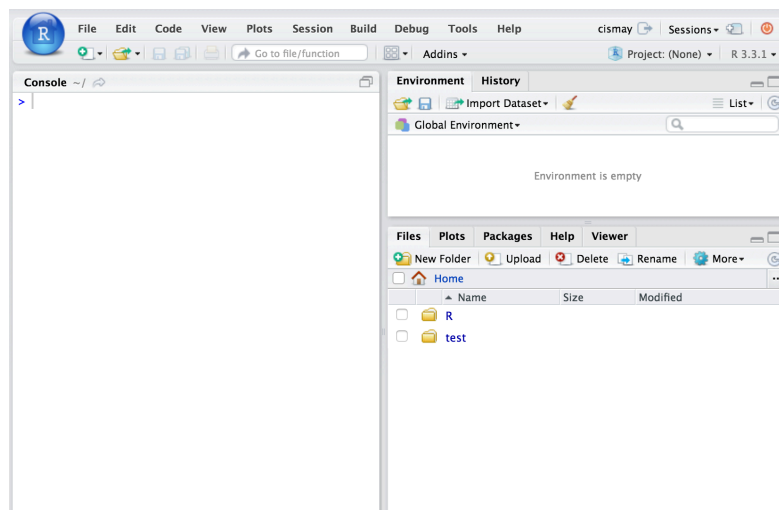
*Starting the labs**Logging in and initial screen*

The RStudio Server provides a web-based way to run analyses in R. This means that you will only need an internet connection and a web browser to run your analyses. You can be running a Windows machine, a Mac, a Linux, or pretty much any other device that has access to the internet and a web browser. If you are interested in using the Reed College RStudio Server off campus, you'll need to request access in the form [here](#). The initial login page will not load for you unless you are connected to the VPN from off campus locations. After entering the link (<http://rstudio-dev.reed.edu>) into your browser, you'll see a page that looks something like:



The image shows the RStudio login interface. At the top is the RStudio logo. Below it is a box titled "Sign in to RStudio". Inside this box, there are two input fields: "Username:" and "Password:". Below the password field is a checkbox labeled "Stay signed in". At the bottom of the box is a blue button labeled "Sign In".

After logging in with your Reed College username (mine is `cismay`, for example) and password, you should see a layout similar to what follows.



Creating and Sharing Projects using RStudio Server

A good habit to get into whenever you start a new project following a common theme (such as your CHEM 101 labs) with R code is to create a new RStudio project to go along with it. RStudio project files have the extension `.Rproj` and store metadata that goes along with the documents you've saved and information about the R environment you are working in. More information about RStudio projects is available from RStudio, Inc. [here](#).

You'll be sharing your Chem 101 labs with your Danielle Cass (`dcass`) and Chester Ismay (`cismay`) using the RStudio Server. This will allow both of us to help you with errors you may be having in writing your lab report without having to email files back and forth. You'll just need to email us with the name of the lab you are having troubles with and what you named your RStudio project and we'll be able to go in and look over your R Markdown file. You can think of this as working somewhat like Google Docs in that both Danielle and I can provide feedback and walk-through your analyses.

To share the project with us, you should follow along with the example below. Instead of naming your RStudio project `CHEM101_ChesterIsmay` as the GIF shows, you should name it `CHEM101_YourName` where `YourName` is replaced with your first and last names. Spaces can cause problems in file names so it is recommended that you avoid them. If you've already shared your project and you have spaces in the name, you'll probably be OK, but try to avoid spaces in your file names going forward. Calling your project `CHEM101_YourName` will allow Danielle and I to track down your project more easily than if everyone named their project `initial` or something similar.

https://raw.githubusercontent.com/ismayc/chemistr-book/master/gifs/proj_share.gif

You'll need to also add `cismay` as someone you'd like to share the project with. You can do so by entering my username immediately after you have clicked **Add** to add Danielle or you can go back to **Share Project** and enter my name there just as you have with `dcass`. If you see that your project is shared with both `dcass` and `cismay`, you have completed this stage successfully.

You'll see at the beginning of each chapter corresponding to the labs how you can access the different R Markdown template files for each lab. These template files are built into the `chemistr` package that I have already loaded on the Reed RStudio Server for you. ¹

Don't overwrite the Rproj file!

Important Note: There is a BIG difference between an Rmd (R Markdown) file that will store the code for each of your labs and this Rproj (RStudio Project) file. **DO NOT OVERWRITE the Rproj file with your Rmd file or any other file.** If you do this, Danielle and I will not be able to assist you since the Rproj file stores important information about sharing projects. You'll need to create a new project all over from the beginning and it's much better for everyone to not overwrite this file. Please create a new folder for each lab and **LEAVE THIS Rproj file ALONE.**

¹ If you are completing the labs using RStudio Desktop, you'll see instructions on how to install this package onto your computer in the Appendix. After installing the package, you'll be able to use the R Markdown templates in much the same way students using the RStudio Server can. Note that if you are using the RStudio Desktop, you won't have access to the Shared Projects feature and it will be more difficult for Danielle and I to help you.

Requesting help

It's nearly impossible for Danielle and I to help you unless we know that you are having troubles. If you find yourself unable to complete a lab because of problems with RStudio and you've spent more than 30 minutes trying to figure it out, it is time for you to stop and ask for help. We encourage you to try to figure it out on your own as much as possible because this helps you learn, but if you are deeply entrenched in not getting R to work, that could be taking away from your learning of Chem 101/102 material.

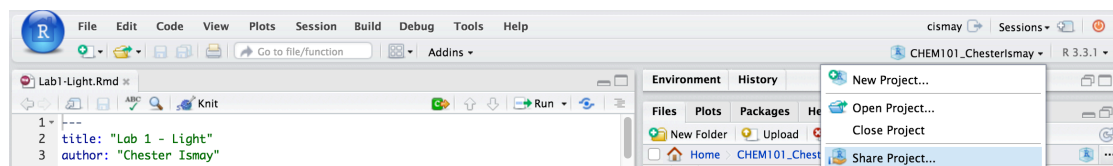
We encourage you to reach out to both of us and carefully explain what is going wrong with your R Markdown file. Sending us an email saying "I can't get it to work, please help." is not insightful and won't help you learn in the long-run. Be as descriptive as possible. Something like what follows is much preferred:

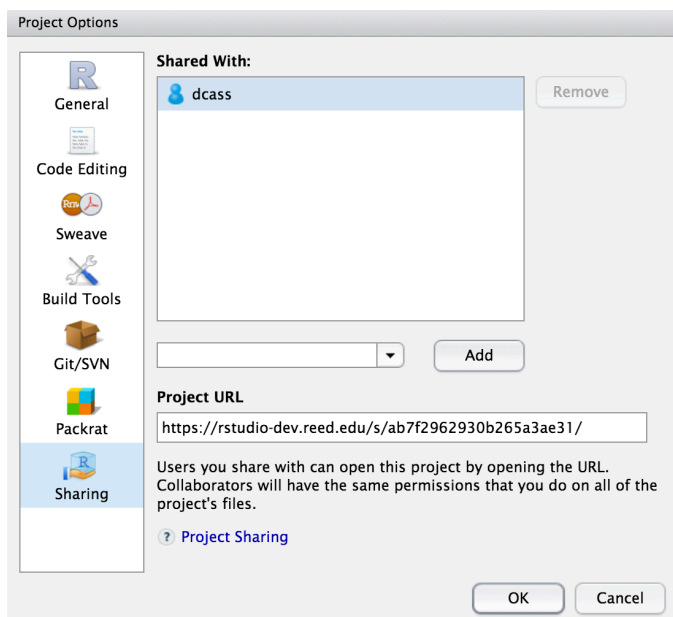
"I'm receiving an error on line 46 of my Light lab about not finding an object. I've tried Googling it and found a couple references about it including [link1](#) and [link2](#) but it's not working for me."

This tells us that you have tried to figure it out for yourself, but that you require our assistance. That is to be expected and you shouldn't be discouraged if you don't understand something about R Markdown immediately.

We are here to help and please don't give up on this. Learning how to work with R even at the most basic of levels will help you long-term in keeping track of analyses and getting you on the path to reproducible research. This will help you in your academic careers and whatever you do after your academic career ends. If you need assistance, please email Danielle and Chester with your concise error specifying in which lab and what the name of your RStudio project is. We are happy to help you learn how to use R!

Additionally, it is extremely helpful for us if you can share a link to your RStudio Project in any emails requesting help. This link is available by going to your RStudio project in the top right corner of RStudio, clicking on it and then selecting **Share Project**, and then select **Sharing** as seen in the screenshots below.





The link is given in the **Project URL**. Please copy this entire link into the body of your emails to Danielle or I so that we can quickly look into your errors.

Book was last updated:

```
## [1] "By dcass on Thursday, June 08, 2017 02:08:56 PM PDT"
```


1

Lab 1: Light

1.1 Getting the lab template

In RStudio, you'll find the R Markdown template that you will update with your own results by following along with the GIF below:

https://raw.githubusercontent.com/ismayc/chemistr-book/master/gifs/light_template.gif

I've also created a new folder called `Lab1-Light` in my `CHEM101_ChesterIsmay` folder and named the resulting R Markdown file that was loaded in from the R Markdown template in `chemistr` the same thing: `Lab1-Light.Rmd`. You are highly encouraged to follow this same convention for all of your labs. (Note that as I mentioned earlier in this manual I do not have any spaces in the folder names or in the names of files here.) This will allow you to remember exactly where you stored the files and allow Danielle and I to easily find your files on the RStudio Server. Note that this folder was also created in the `CHEM101_ChesterIsmay` folder where the `Rproj` file resides. If you put your lab in the `Home` directory instead, Danielle and I won't be able to access it since `Home` is the parent directory of `CHEM101_YourName`.

1.2 The parts of the lab template

1.2.1 The YAML header

I've already entered some values into my R Markdown file. First I gave a title to the R Markdown file: "Lab 1 - Light". I then entered my name and the current date. All of these are in quotation marks since they all correspond to character strings and not numbers. It will likely still work if you don't put quotation marks around the entries here, but it is recommended that you use the quotation marks to help you understand the use of quotation marks when you begin working with R code. (R won't be as nice as YAML is on forgetting quotation marks when they are needed.)

This portion of the R Markdown file is called the YAML header. It stores metadata about your document. It is separated by `---` before and `---` after on their own lines. RStudio will warn you by coloring the `---` red if you accidentally add a space before them.

You'll need to make sure to include a space after each of the YAML entries. You'll receive an error if you have something like that below, for example:

```
title: "Light Lab"
author: "Chester Ismay"
date:"August 30, 2016"
output: chemistr::chem_lab_word
```

There needs to be a space after the `:` for the `date` field.

The last line here corresponds to the type of output that will be produced. In this case, I've created a nice way for you to output this R Markdown file to a DOCX (Word document) format. The `chem_lab_word` document format in `chemistr` creates a Word document with figure captions enabled by default and images with size 3 inches in height and 6 inches in width by default.

1.2.2 Initial R Chunk

On line 8 of the Light R Markdown template, you'll see three backticks starting the line. This tells R Markdown to expect a block/chunk of code. This code is R code by the next designation `{r}`. Next we give this R chunk a name `setup` and then specify chunk options immediately after the name and a comma. Here we have specified `include=FALSE`. We then close the specifications for the chunk with a `}`. This `include=FALSE` chunk option will run the R code when we press **Knit** button near the top of the code editor will not put the R code or its output into the document. We'll press the **Knit** button in a bit to understand what it accomplishes.

You'll see this `setup` chunk in all of the lab templates. It's good practice to load any package you'll need for your analysis at the top in a chunk like this. You may not initially understand why `chemistr` is needed but it will be needed when we create a table using the `chem_table` function, which resides in the `chemistr` package, in the `table` chunk beginning on line 21.

1.2.3 Headers

Immediately below this first R chunk is two hash tags. This corresponds to creating a header in our resulting document. The font immediately following the hash-tags will be larger and act as a divider in our document. The number of hash-tags you enter corresponds to how large you'd like to make the header. One hash-tag is the largest and six is the smallest. It is recommended that you use between one and three hash-tags for headers.

1.2.4 Enter data chunk

If you press the **Knit** button at this moment, you'll receive an error. Let's look into what this error means:

https://raw.githubusercontent.com/ismayc/chemistr-book/master/gifs/knit_error1.gif

Danielle has created lab templates that might not necessarily build without errors immediately for you. You should carefully read over the directions in each chunk (The comments after the `#` in the R chunks.) before pressing **Knit**. Here we see that the error occurs near line 15 of our R Markdown file. The `Color` and `WavelengthRange` vectors need to have some values entered. These results will correspond to what you observed in lab. I will enter some demo results in here and also show you how to use the R Console to check your results. You'll see that the `Color` and `WavelengthRange` vectors must be of the same length when we create a `data_frame` to combine them together into "spreadsheet" format in the variable named `Q1_data`:

https://raw.githubusercontent.com/ismayc/chemistr-book/master/gifs/enter_data_full.gif

You should use the R Console to check your results in the R chunks by either entering the R code line by line into the Console or by pressing the **Run Current Chunk** button (the green Play button). We see here that `Q1_data` is 6 rows by 2 columns and stores our results in "spreadsheet" form as a `data_frame`.

Note that you'll also receive an error if you have a different number of entries for `WavelengthRange` and `Color`. Data frames must have an equal number of rows when you combine the columns together via `data_frame`.

The `chemistr` package also loads the `dplyr` package where the `data_frame` function resides. If you are running the chunks by pressing the green "play" button, which sends the current chunk into the R console, you'll need to remember to load all previous chunks as well. If you receive an error message such as

```
Error: could not find function "data_frame"
```

this means that you have not loaded the `chemistr` package by sending the `library(chemistr)` code into the Console BEFORE you run the declarations for `Color`, `Wavelength`, and `Q1_data <- data_frame(Color, WavelengthRange)`. The code chunks read like a book so if you are receiving errors of the "could not find" variety, it probably means you haven't run all chunks up to the point of your current chunk in the Console. These sorts of errors are also referenced in Chapter 6 of "Getting Used to R, RStudio, and R Markdown" (Ismay, 2016).

1.2.5 A nice table R chunk

The `chem_table` function will result in a nicely formatted table being produced in your resulting Word document. You'll need to specify a caption inside the parentheses instead of the "A lovely caption for the table" default. You'll also see that RStudio will create a second quotation mark when you might not want one. Carefully look over your document if you have errors and make sure there aren't two quotation marks when you only wanted one. You'll see this below:

https://raw.githubusercontent.com/ismayc/chemistr-book/master/gifs/two_quotes.gif

After fixing our error, we have now created a DOCX file that will have our nice table showing the `Q1_data`. Feel free to open up this `Lab1_Light.docx` file now and see that the table was created with the caption "Light Source Table". You should probably enter a different and more

specific caption instead though.

If you scroll down the R Markdown file, you'll see that there are many more spots where we will be inserting pictures. R Markdown has nicely skipped over these chunks when we pressed **Knit** instead of halting and giving lots of confusing errors and we see that in the R Console window under the R Markdown tab.

1.3 *Uploading and inserting external pictures*

You'll frequently be asked to add pictures into your lab reports that weren't necessarily created using R. (We will see how to include plots created inside R in later labs though.) If you look over **Q2** in the lab it is asking you to include a picture of the "Absorbance spectrum of water in a 1 cm cuvette". Note that this is specified as the figure caption here for the first R chunk after **Q2**.

Mathematical text is also included by surrounding the mathematical expression with dollar signs. This is using LaTeX to produce nicely formatted math text. Here we see the expression H_2O is given using `$H_{2}O$`.

In order to include a picture, you'll need to first upload it to the RStudio Server into the directory where the R Markdown file is located. The GIF below shows you how to do so. Remember that you'll need to know where you saved your image on your computer to specify it here. This example GIF is recorded on a Mac but similar procedures should be done on a Linux or Windows machine.

https://raw.githubusercontent.com/ismayc/chemistr-book/master/gifs/upload_photo.gif

The `Water.jpg` file has been uploaded to my `Lab1-Light` folder on the RStudio Server. The `include_graphics` function requires that you add the name of the file and its extension here. We now press the **Knit** button to see the resulting Word document:

<https://raw.githubusercontent.com/ismayc/chemistr-book/master/gifs/knitted.gif>

1.3.1 *The remainder of the lab*

In the remaining portions of the lab you'll be asked to include other pictures from your lab sessions and also include a discussion of your results. This discussion will proceed in much the same way as you would add a discussion into a Word document or Google Doc. You'll be able to immediately discuss what occurred in your R chunks right below them. This allows for the reader to easily follow your work.

1.4 *Note on white space*

As you look over the R Markdown document you'll see that there is always a new line of white space between the discussion and the R chunks and also between each of the R chunks. It is highly recommended that you also follow this workflow. You'll receive some strange errors at

times if you try to stack everything together and it's also much harder to follow for another reader of your document if you have everything bunched together. **White space is your friend!**

1.5 *Spell-check*

Just as I'm sure your English teachers have told you to spell check your documents before submitting, you are also encouraged to do so here. There is a built-in spell check option found near the **Knit** button.



Please run this and carefully read over your lab report before converting it to a PDF and submitting it to Moodle.

1.6 *Converting your Word document to PDF*

The directions for each lab on Moodle say to upload a PDF version of your lab. You'll see how to create this PDF from inside Microsoft Word for Mac. A similar procedure can be done using LibreOffice (<https://www.libreoffice.org/download/libreoffice-fresh/>) on Linux, Mac, or Windows machines or Microsoft Word on a PC. You may also have the option to **Save As** a PDF there and you can get to this option by going to **File -> Save As -> File Format: -> PDF** on the Mac if you prefer.

https://raw.githubusercontent.com/ismayc/chemistr-book/master/gifs/word_pdf.gif

2

Lab 2: Beer's Law

2.1 Getting the R Markdown lab template

Begin by creating a new folder called **Lab2-Beers** in your **CHEM101_FirstnameLastname** folder. Next, follow steps similar to those found in the GIF in Section 1.1, but select the **Beer's Law** R Markdown template instead of **Light** and save your file in the newly created **Lab2-Beers** folder as **lab2.Rmd**.

2.2 The parts of the lab template

2.2.1 YAML

Refer back to The YAML header section in Chapter 1 for a review on what the entries here mean. Remember to be careful with spacing!

2.2.2 Initial R Chunk

Remember that the **chemistr** package automatically loads in many useful packages for you. It also includes functions such as **chem_table** you worked with in Lab 1 and **chem_scatter** that you will work with in this lab.

2.2.3 Directions chunk

Immediately after the **## Data** header, you'll see a chunk of commented R code Danielle has including giving you directions on what the chunks after it should include. Note that this chunk again uses **include=FALSE** so you won't see any of this in the resulting knitted DOCX file.

2.2.4 First data chunk

After carefully reading over the commented R code Danielle has added providing you with directions, you now are ready to enter values in for Part A. Note that you can change the vari-

able names from `Independent1` and `Dependent1` to something else, but remember that those names cannot include spaces or other mathematical objects. Remember that if you change the name of `Independent1` to, say, `my_indep` that you'll need to make sure to change all references from `Independent1` to `my_indep` in the R code that follows.

2.2.5 Transforming to a logarithmic scale

We often need to look at data not on our usual scale but instead on a logarithmic scale. With logarithms we specify a base and you may remember from algebra that logarithms are related to exponential functions. To further emphasize this, recall that $10^2 = 100$ can be written as $\log_{10} 100 = 2$.

It is important to note that $\log_{10} 0 = -\infty$ so you'll need to be careful with how you handle values of 0 when you calculating a logarithmic function on a vector including 0. For this lab, you are encouraged to remove any 0 values from your vectors BEFORE you take the logarithm to avoid this problem.

In this lab, you'll be using "log base 10" and the `log10` function in R. I recommend you enter `?log10` into the R Console for more information on how to do a base 10 logarithmic transformation on a vector in R. Additionally, you can find a great explanation of how logarithms are used in the real world here.

One of the powerful features of R is its ability to do vectorized calculations. Here is an example:

```
my_nums <- c(10, 15, 105, 500, 100231)
log10(my_nums)
```

```
## [1] 1.000000 1.176091 2.021189 2.698970 5.001002
```

The code above produces the "log base 10" for each of the values in `my_nums`. We haven't assigned this to anything though since we didn't include the assignment operator `<-`. In the GIF below, I'll show you how to create one of the two new variables you'll need to create. Remember that you'll also need to update your `data_frame` call to include both of your newly created variables as well.

<https://raw.githubusercontent.com/ismayc/chemistr-book/master/gifs/log.gif>

As I have in the GIF, you should use the R Console to check your results in the R chunks by either entering the R code line by line into the Console or by pressing the **Run Current Chunk** button (the green Play button).

2.2.6 Copying from one chunk into a new chunk

When you are first learning to program, one of the best strategies you can do is to copy working code and make small changes to the code to produce a different result. In the `plot1` chunk (remember to look at the phrase right after the `{r` and before the comma that begins your R chunk to find the name), you will see how to produce a scatterplot of `Dependent1` on the vertical axis and `Independent1` on the horizontal axis (frequently written as "a scatterplot of

Dependent1 versus Independent1—or y versus x). In the GIF below, you’ll see how we could produce a plot of the log base 10 of Dependent1 versus Independent1. Notice that only a few subtle changes are needed to produce a different plot. Also think about how the new plot in chunk plot2 compares to the plot produced in plot1.

<https://raw.githubusercontent.com/ismayc/chemistr-book/master/gifs/log2.gif>

You’ll need to use keyboard shortcuts to copy-and-paste code. (Ctrl + C on a Windows machine or Command + C on a Mac to copy. Then use Ctrl + V on a Windows machine or Command + V on a Mac to paste.)

2.2.7 Regression

Carefully note in the GIF how I have updated the variables to produce a new fit and a new plot. The `lm` function expects `y ~ x` so make sure you have them in the correct order. In other words, you’ll want to have your “dependent variable” ~ “independent variable”. This is used to get the coefficients from a straight line regression fit of the data. The `summary` function produces a lot of important information about this fit.

2.2.8 The remainder of the lab

In the remaining portions of the lab, you are asked to modify the results above to make 4 additional plots. Additionally, you’ll need to provide commentary text below the Discussion header.

2.3 Note on white space

As you look over the R Markdown document you’ll see that there is always a new line of white space between the discussion and the R chunks and also between each of the R chunks. It is highly recommended that you also follow this workflow. You’ll receive some strange errors at times if you try to stack everything together and it’s also much harder to follow for another reader of your document if you have everything bunched together. **White space is your friend!**

2.4 Spell-check

Just as I’m sure your English teachers have told you to spell check your documents before submitting, you are also encouraged to do so here. There is a built-in spell check option found near the **Knit** button.



Please run this and carefully read over your lab report before converting it to a PDF and submitting it to Moodle.

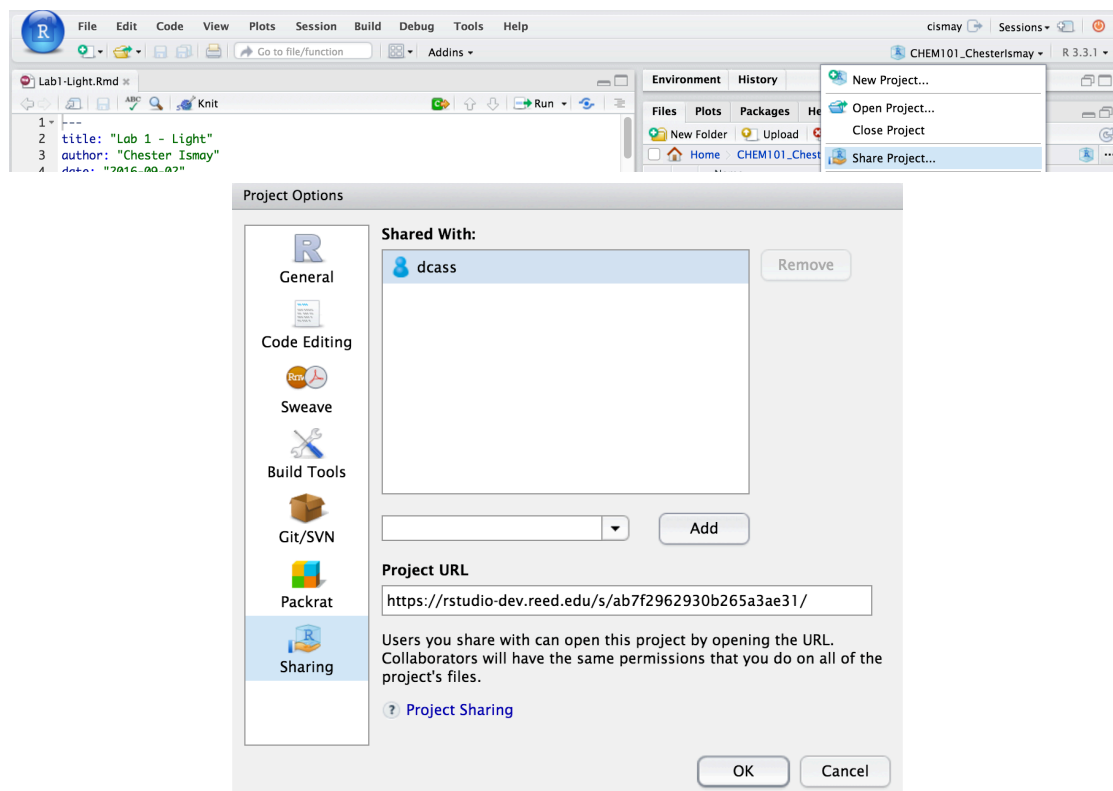
2.5 Converting your Word document to PDF

The directions for each lab on Moodle say to upload a PDF version of your lab. You'll see how to create this PDF from inside Microsoft Word for Mac. A similar procedure can be done using LibreOffice (<https://www.libreoffice.org/download/libreoffice-fresh/>) on Linux, Mac, or Windows machines or Microsoft Word on a PC. You may also have the option to **Save As** a PDF there and you can get to this option by going to **File -> Save As -> File Format: -> PDF** on the Mac if you prefer.

https://raw.githubusercontent.com/ismayc/chemistr-book/master/gifs/word_pdf.gif

2.6 Note on requesting help

It is extremely helpful for us if you can share a link to your RStudio Project in any emails requesting help. This link is available by going to your RStudio project in the top right corner of RStudio, clicking on it and then selecting **Share Project**, and then select **Sharing** as seen in the screenshots below.



The link is given in the **Project URL**. Please copy this entire link into the body of your emails to Danielle or I so that we can quickly look into your errors.

3

Lab 3: Reactor

3.1 Getting the R Markdown lab template

Begin by creating a new folder called **Lab3-Reactor** in your **CHEM101_FirstnameLastname** folder. Next, follow steps similar to those found in the GIF in Section 1.1, but select the **Reactor** R Markdown template instead of **Light** and save your file in the newly created **Lab3-Reactor** folder as **lab3.Rmd**.

3.2 The parts of the lab template

3.2.1 YAML

Refer back to The YAML header section in Chapter 1 for a review on what the entries here mean. Remember to be careful with spacing!

3.2.2 Initial R Chunk

Remember that the **chemistr** package automatically loads in many useful packages for you. It is loaded here in the **setup** chunk and will need to be loaded in order to complete the plots in this lab. It also includes functions such as **chem_table** you worked with in Lab 1 and **chem_clustered.bar** and **chem_dual.scatter** that you will work with in this lab.

3.2.3 Results chunk

Immediately after the **## Results** header, you'll see a chunk of code with name **chunk1** corresponding to your data from part A. You will have three values for each of the four different shields you used. Thus, the **material** variable will contain each of the shields 3 times using the **c** function which combines objects together and the **rep** function which repeats values a specified number of **times**. For each shield you'll have a corresponding element value for each of the three elements. Thus, the **element** vector is simply the three elements repeated 4 **times**. Lastly, you'll need to enter the %T values in the **value** variable. The last line creates a **data_frame** that has 12 rows and 3 columns. You are encouraged to enter **View(shield_data)**

in the R console AFTER you have run this chunk and the previous chunk. This will give you a glimpse into the table layout of the data frame and also allow you to make changes to your `value` entry as needed.

3.2.4 *Clustered bar graph chunk*

In `chunk2`, you will be producing a bar graph that allows you to look at how the different `element` values relate with the different shielding values in terms of %T. You also have the option to change the colors you'd like to use by entering the names in quotation marks in a `named_colors` vector that is passed into the `chem_clustered.bar` function via the `colors` argument. As noted in the R comments in `chunk2`, you can find a listing of all of the named colors in R at <http://www.stat.columbia.edu/~tzheng/files/Rcolor.pdf>.

In the GIF below, I walk through the recommended steps to download and save the template file in the appropriate location in your folder on the RStudio Server. Additionally, you'll see what each of the different variables correspond to in the `shield_data` data frame and how they are plotted in the clustered bar graph.

https://raw.githubusercontent.com/ismayc/chemistr-book/master/gifs/cluster_bar.gif

As I have in the GIF, you should use the R Console to check your results in the R chunks by either entering the R code line by line into the Console or by pressing the **Run Current Chunk** button (the green Play button). Additionally, you are encouraged to run all previous chunks before running the current chunk by pressing the button just to the left of the green Play button.

3.2.5 *Part B data entry*

In `chunk3` you'll enter data in much the same way as you did in previous labs. Here you'll specify different values for `thickness`, `abs`, and `trans`. Note here that these are all numeric values and so you don't (AND SHOULDN'T) enter them with quotes around them.

3.2.6 *Dual scatter plot*

In `chunk4` you are presented with code to plot two vertical axis numerical variables to go with one horizontal numerical variable. Here we specify `thickness` as the horizontal and `abs` and `trans` as the vertical axes. This should be an extension of the work done in Lab 2 when you produced a scatterplot via the `chem_scatter` function.

3.2.7 *The remainder of the lab*

Lastly, you'll need to provide commentary text below the Discussion header following the Exp 3 lab instructions.

3.3 Note on white space

As you look over the R Markdown document you'll see that there is always a new line of white space between the discussion and the R chunks and also between each of the R chunks. It is highly recommended that you also follow this workflow. You'll receive some strange errors at times if you try to stack everything together and it's also much harder to follow for another reader of your document if you have everything bunched together. **White space is your friend!**

3.4 Spell-check

Just as I'm sure your English teachers have told you to spell check your documents before submitting, you are also encouraged to do so here. There is a built-in spell check option found near the **Knit** button.



Please run this and carefully read over your lab report before converting it to a PDF and submitting it to Moodle.

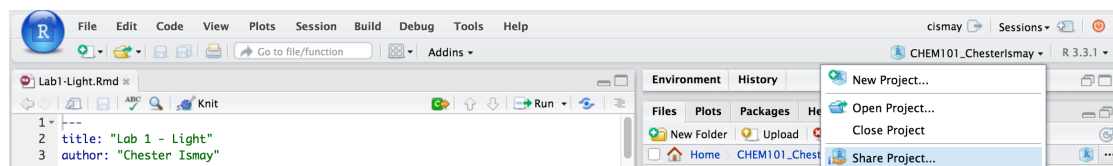
3.5 Converting your Word document to PDF

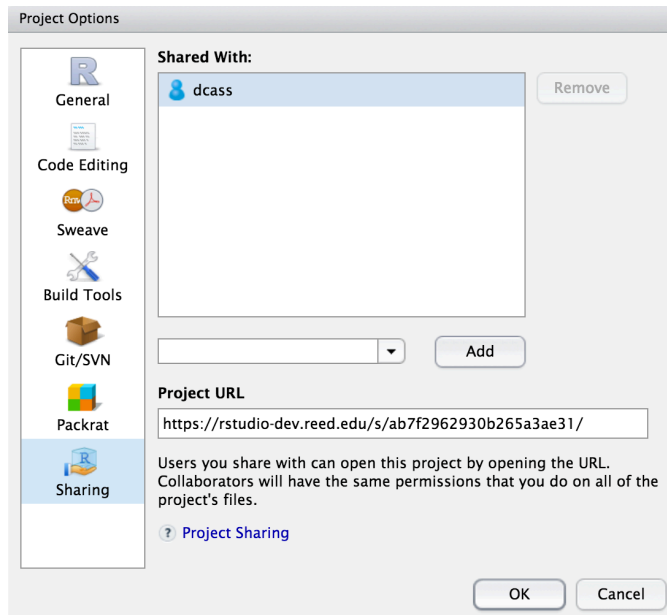
The directions for each lab on Moodle say to upload a PDF version of your lab. You'll see how to create this PDF from inside Microsoft Word for Mac. A similar procedure can be done using LibreOffice (<https://www.libreoffice.org/download/libreoffice-fresh/>) on Linux, Mac, or Windows machines or Microsoft Word on a PC. You may also have the option to **Save As** a PDF there and you can get to this option by going to **File -> Save As -> File Format: -> PDF** on the Mac if you prefer.

https://raw.githubusercontent.com/ismayc/chemistr-book/master/gifs/word_pdf.gif

3.6 Note on requesting help

It is extremely helpful for us if you can share a link to your RStudio Project in any emails requesting help. This link is available by going to your RStudio project in the top right corner of RStudio, clicking on it and then selecting **Share Project**, and then select **Sharing** as seen in the screenshots below.





The link is given in the **Project URL**. Please copy this entire link into the body of your emails to Danielle or I so that we can quickly look into your errors.

4

Lab 4: Iron

4.1 Getting the R Markdown lab template

Begin by creating a new folder called **Lab4-Iron** in your **CHEM101_FirstnameLastname** folder. Next, follow steps similar to those found in the GIF in Section 1.1, but select the **Iron Analysis** R Markdown template and save your file in the newly created **Lab4-Iron** folder as **lab4.Rmd**.

4.2 The parts of the lab template

4.2.1 YAML

Refer back to The YAML header section in Chapter 1 for a review on what the entries here mean. Remember to be careful with spacing!

4.2.2 Initial R Chunk

Remember that the **chemistr** package automatically loads in many useful packages for you. It is loaded here in the **setup** chunk and will need to be loaded in order to complete the plots in this lab.

4.2.3 Results chunk

As with previous labs, here you'll enter the values for **Iron** and **Absorbance** and create a data frame named **iron_data** that pairs these two vectors together. Remember that **Iron** and **Absorbance** must be of the same length (have the same number of values) for the **data_frame** function to work.

4.2.4 Plot chunk

We are interested in fitting a linear regression line using **Iron** as our predictor variable and **Absorbance** as our response variable here. Make sure to add labels to your plots and include

an appropriate figure caption in the R chunk option `fig.cap` following the guidelines in the commented R code in the `plot1` chunk.

4.2.5 Calculations

In this block of code you'll be doing some statistical calculations. If you'd like to use R to do so, you'll need to first download the data file from Moodle and read it into an object in R. So, for example, if the data was saved as a CSV file called `iron_class.csv`, you could use the following code after you have Uploaded the file onto your directory on the RStudio Server.

```
iron_class <- read.csv("iron_class.csv")
```

Now that you have the data read in, you can use the `group_by`, `summarize`, and `mutate` functions in R to calculate averages, standard deviations, and confidence intervals. If the name of our iron variable for the class is `iron`, we can use the `summarize` function to calculate the **mean** (average), **standard deviation**, and **n** (how many values are in our sample). We can then use the `mutate` function to create new columns corresponding to the lower value of the confidence interval and the upper value of the confidence interval:

```
iron_class %>%
  summarize(mean.iron = mean(iron, na.rm = TRUE),
            sd.iron = sd(iron, na.rm = TRUE),
            n.iron = n()) %>%
  mutate(se.iron = sd.iron / sqrt(n.iron),
         lower.ci.iron = mean.iron - qt(1 - (0.1 / 2), n.iron - 1) * se.iron,
         upper.ci.iron = mean.iron + qt(1 - (0.1 / 2), n.iron - 1) * se.iron)
```

To review, the `iron_class` data frame is passed into the `summarize` function, which creates three new variables `mean.iron`, `sd.iron`, and `n.iron`. We need to determine `n.iron` in order to calculate the standard error denoted as `se.iron`. We can then put it all together to calculate the confidence interval. Note here that instead of looking up the values in the t table we can have R look them up for us at the 90% level (this corresponds to the 95th percentile value in the t distribution with `n.iron - 1` degrees of freedom).

4.2.6 The remainder of the lab

Lastly, you'll need to include a picture of your calculations if you didn't use R and also a discussion of your results following the directions in the template.

4.3 Note on white space

As you look over the R Markdown document you'll see that there is always a new line of white space between the discussion and the R chunks and also between each of the R chunks. It is highly recommended that you also follow this workflow. You'll receive some strange errors at times if you try to stack everything together and it's also much harder to follow for another reader of your document if you have everything bunched together. **White space is your**

friend!

4.4 *Spell-check*

Just as I'm sure your English teachers have told you to spell check your documents before submitting, you are also encouraged to do so here. There is a built-in spell check option found near the **Knit** button.



Please run this and carefully read over your lab report before converting it to a PDF and submitting it to Moodle.

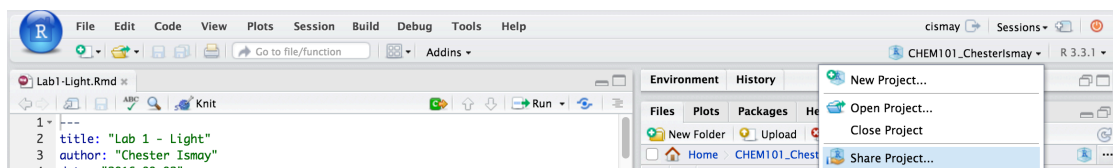
4.5 *Converting your Word document to PDF*

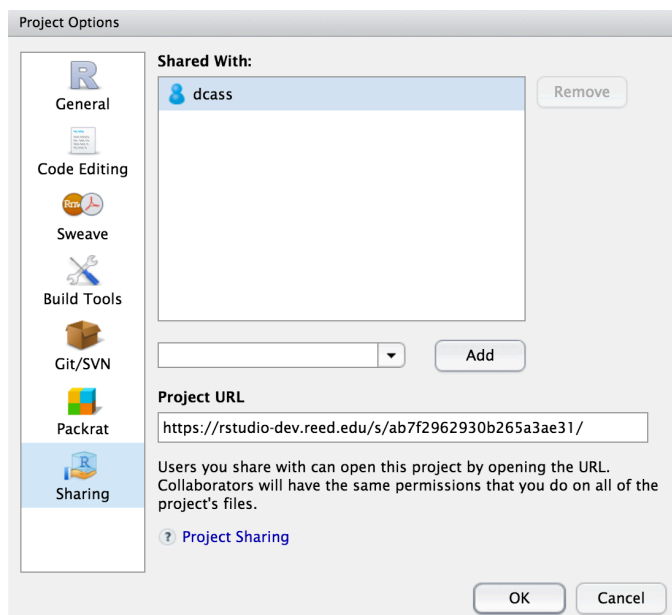
The directions for each lab on Moodle say to upload a PDF version of your lab. You'll see how to create this PDF from inside Microsoft Word for Mac. A similar procedure can be done using LibreOffice (<https://www.libreoffice.org/download/libreoffice-fresh/>) on Linux, Mac, or Windows machines or Microsoft Word on a PC. You may also have the option to **Save As** a PDF there and you can get to this option by going to **File -> Save As -> File Format: -> PDF** on the Mac if you prefer.

https://raw.githubusercontent.com/ismayc/chemistr-book/master/gifs/word_pdf.gif

4.6 *Note on requesting help*

It is extremely helpful for us if you can share a link to your RStudio Project in any emails requesting help. This link is available by going to your RStudio project in the top right corner of RStudio, clicking on it and then selecting **Share Project**, and then select **Sharing** as seen in the screenshots below.





The link is given in the **Project URL**. Please copy this entire link into the body of your emails to Danielle or I so that we can quickly look into your errors.

A

RStudio Desktop

More advanced users of RStudio may be interested in using the RStudio Desktop version installed on their own computer instead of using the RStudio Server. It is encouraged that all students do this after they have some familiarity with how RStudio works. The RStudio Server is a great way to learn how this works while providing the ability for more advanced users to give support to learning through Shared Projects. Downloading the RStudio Desktop allows for resources to be freed up on the RStudio Server though and if you are performing more advanced calculations it makes more sense to download your own version of RStudio instead of running them on the RStudio Server.

A.1 Downloading R and RStudio Desktop

It is worth noting that you can't just install RStudio Desktop without installing R as RStudio needs to have R installed in order to run. A step-by-step guide to installing R and RStudio Desktop with screenshots can be found

- [here](#) for the Mac and
- [here](#) for a PC.

The process is a little more complicated for Linux machines, but Googling “RStudio for Linux” will likely lead you to some instructions. Unless you plan to create PDF directly from R Markdown documents (which requires a multiple gigabyte download of LaTeX) you can skip some of the later steps of the installation guidelines.

A.2 Downloading and installing chemistr

I created the **chemistr** R package as a way to introduce students of Chem 101/102 at Reed College to R without the sometimes intimidating amounts of code needed to produce specific plots and tables. The package also includes R Markdown lab report templates for each of the labs.

Intermediate users of R and RStudio may be used to using the **install.packages** function to download and install R packages from CRAN (The Comprehensive R Archive Network). This network is the standard for R packages and, to be on CRAN, packages there go through a

series of tests to make sure they are working well. With the `chemistr` package being developed recently (and still under development), it does not currently reside on CRAN. But you can still download a developmental version of it onto your machine running RStudio Desktop.

Hadley Wickham has created an R package called `devtools` that allows for R packages to be more easily created and shared with others. We will use this package to install the `chemistr` package from my GitHub page. You'll need to enter these commands into the R Console of your RStudio Desktop:

```
install.packages("devtools")
devtools::install_github("ismayc/chemistr")
```

After running these two commands you can check that `chemistr` was installed correctly by entering `library(chemistr)` into the R Console. If you get back to `>` without any error messages, you should be good to go. Now you'll follow the same steps given in each chapter of this book to get the R Markdown template for that specific lab.

File -> New File -> R Markdown -> From Template -> NameOfLab

You are encouraged to run the two lines of code above before you begin each lab as there may have been slight corrections/changes made to the lab templates. You can find a description of each change to the lab templates here with dates given.

B

Lab 1: Light (More Details)

B.1 The chem_table function

The `chem_table` function is essentially a wrapper function to the `pandoc.table` function in the `pander` package. Here is the code for `chem_table`:

```
chem_table <- function(data, caption){  
  names(data) <- pandoc.strong.return(names(data))  
  pandoc.table(data, caption = caption, style = "multiline",  
               split.tables = Inf)  
  cat("\n\n")  
}
```

We see here that `chem_table` expects two arguments:

- `data` is a data frame containing the variables you'd like to appear in the table
- `caption` is the caption we'd like to correspond to this table

Remember that you can run `?chemistr::chem_table` to bring up the help documentation for the function.

We first use the `pandoc.strong.return` function that bolds the column names. We then call the `pandoc.table` function in the `pander` package with `data` as our argument and then our entered `caption` parameter as the `caption` argument to `pandoc.table`. The last two arguments for `style` and `split.tables` are used to ensure the outputted table appears as you might expect it to in the Word document:

- `style` set to `multiline` allow headers and table rows to span multiple lines of text. This may be helpful if you have long column names.
- `split.tables` set to `Inf` ensures that wide tables will not be split into multiple tables.

Lastly, the `cat("\n\n")` code specifies that an extra line of white space will be printed immediately following the table.

B.2 The *include_graphics* function

The `include_graphics` function enables you to include pictures that are stored as image files (*.png or *.jpg, for example) into your Word document via R Markdown. It is a function in the `knitr` package. If you'd like more information on how to use `include_graphics` run `?knitr::include_graphics` in the R Console. The important argument here is `path` which tells R where to look for the file you want to include.

If you include your pictures in the same directory as your Rmd file you need only specify the name of the file here in quotation marks. If you have a `figure` folder in the same folder as your R Markdown file, you'll need to specify that via something like:

```
knitr::include_graphics("figure/myimage.png")
```

Whatever you specify as the chunk option `fig.cap` will appear as the figure caption. You are also encouraged to name the R chunk, which will allow you to reference the figure later in your document as well as automatic numbering of the figures:

```
```${r myimage, echo=FALSE, fig.cap="Here is my picture"}
knitr::include_graphics("figure/myimage.png")
```
```

We could then reference our picture in the text of our document by using `\@ref(fig:myimage)`.

References

Ismay, C. (2016). *Getting used to R, RStudio, and R Markdown*.