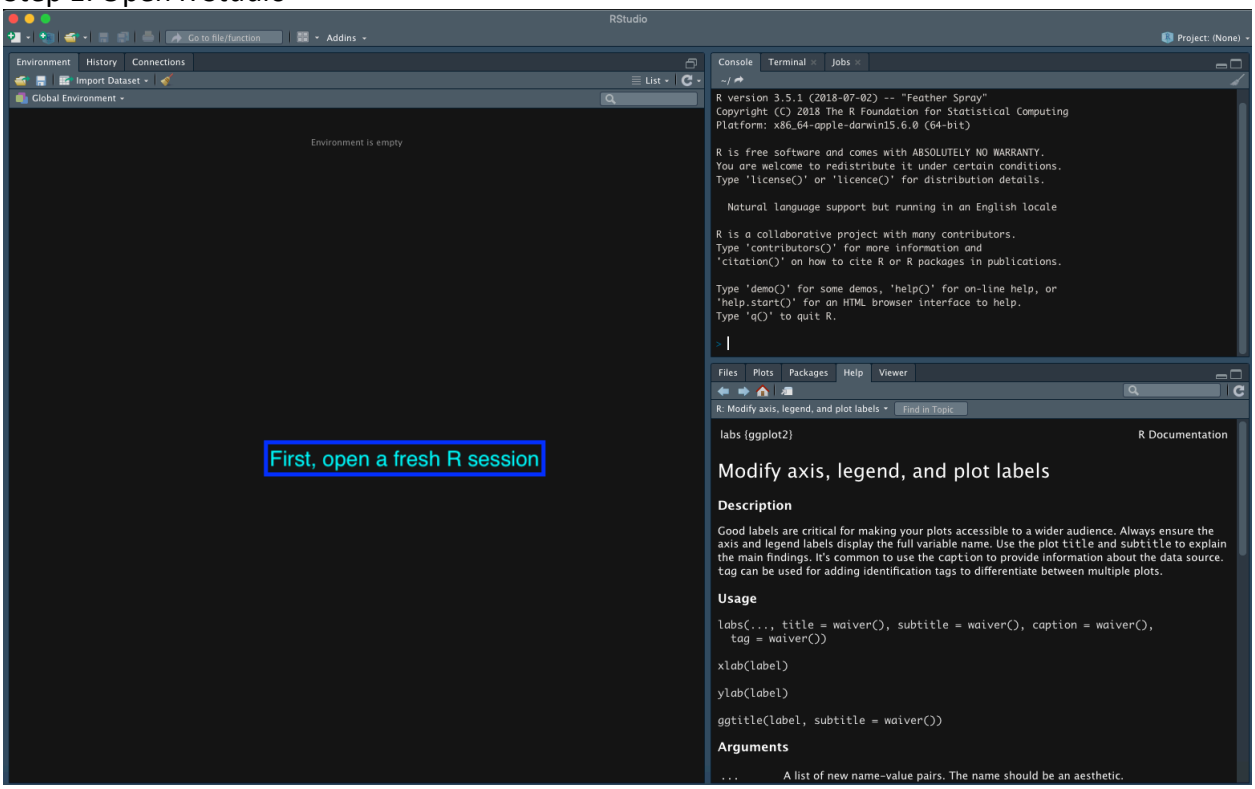
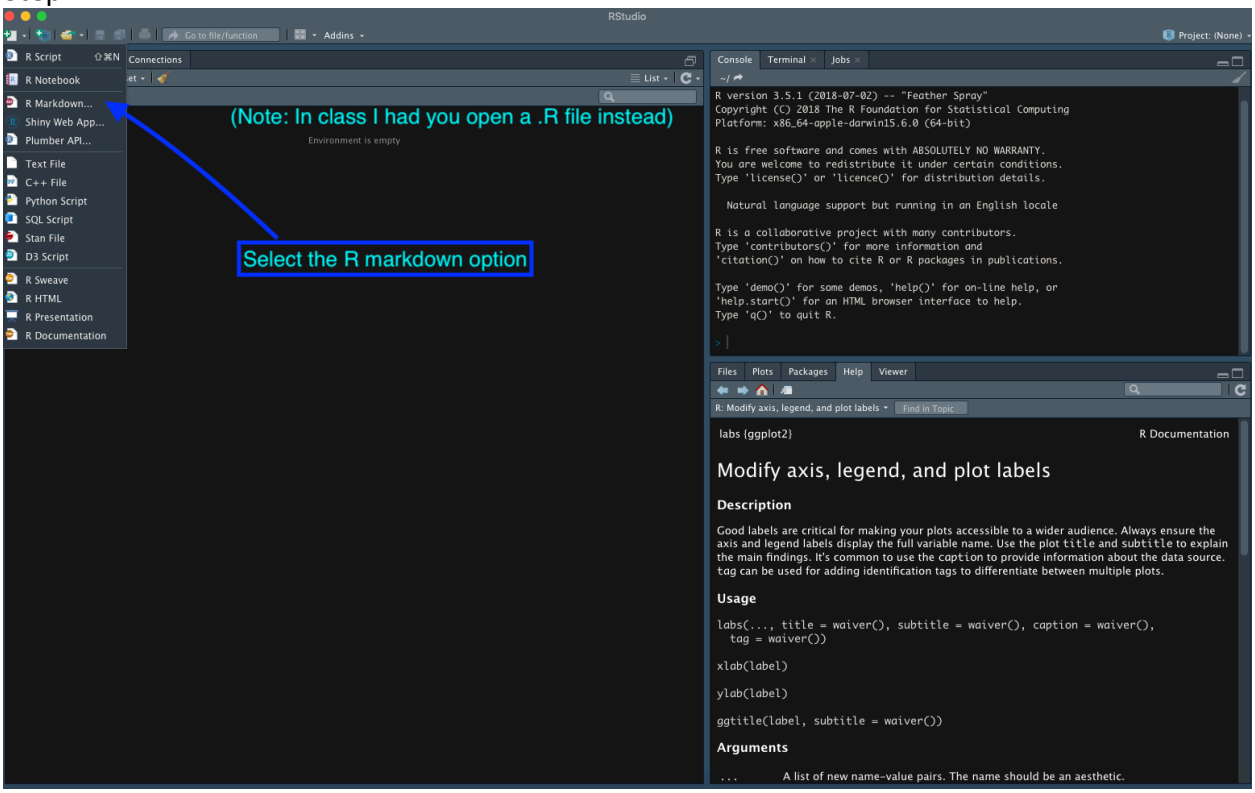
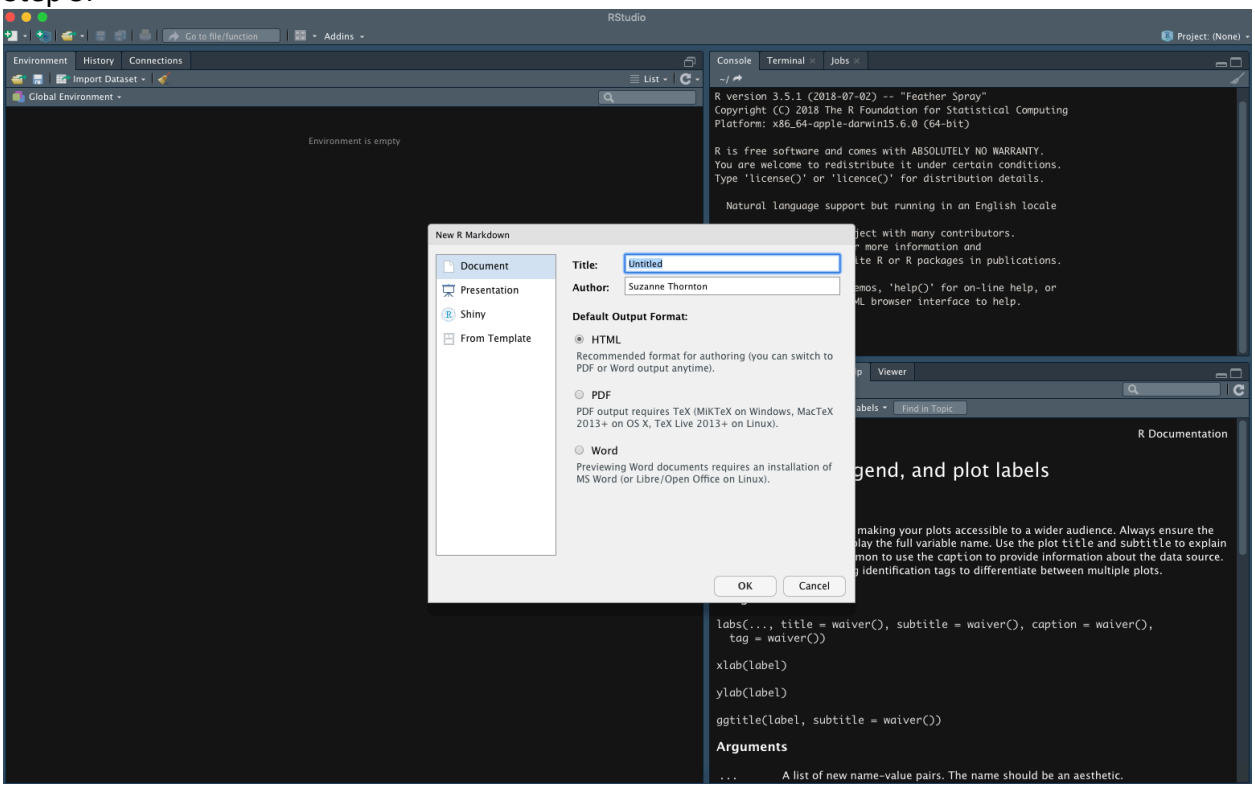
Step 1: Open R Studio



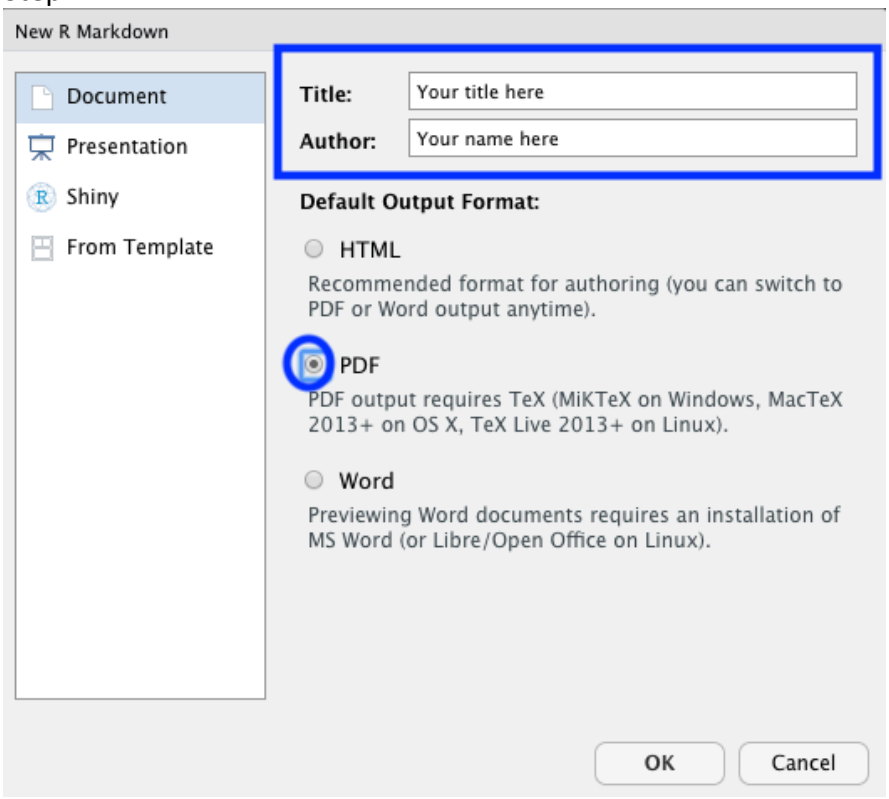
Step 2:



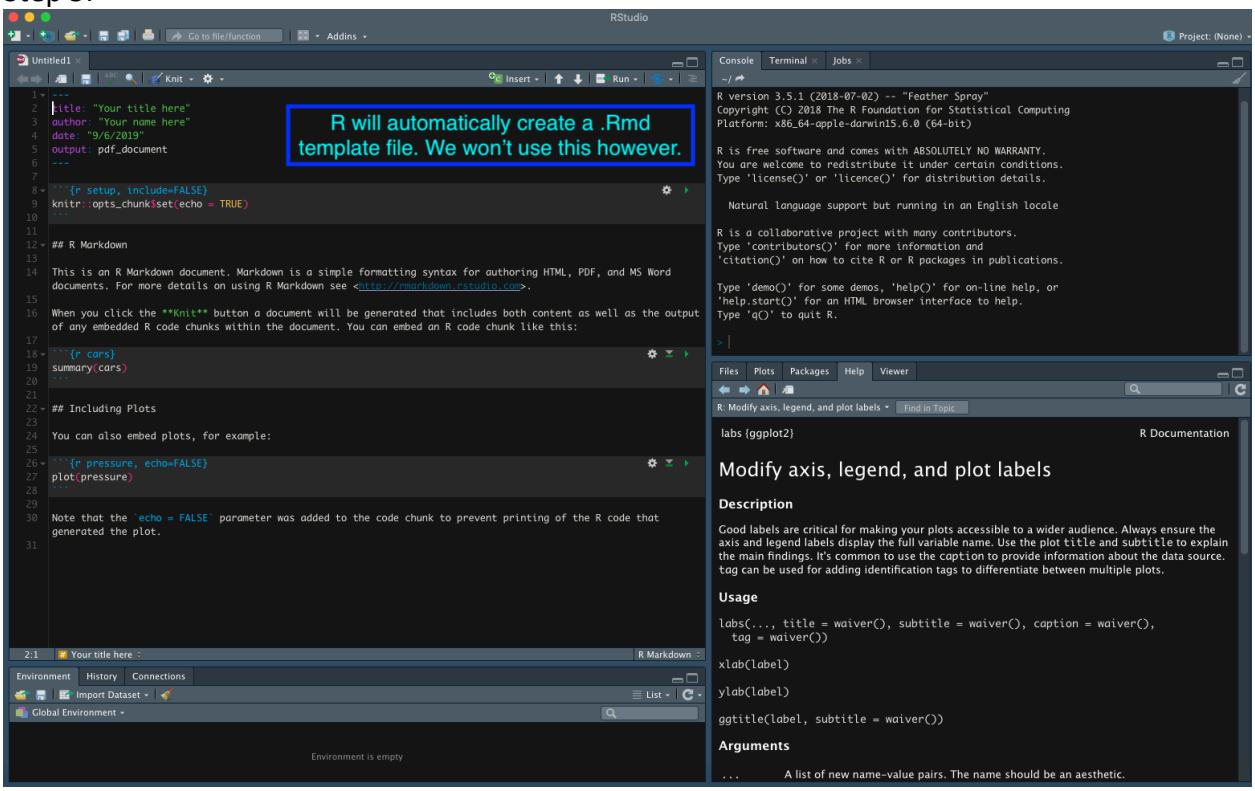
Step 3:


RStudio
Environment
History
Connections
Import Dataset
Global Environment
Environment is empty
Console
Terminal
Jobs
R version 3.5.1 (2018-07-02) -- "Feather Spray"
Copyright (C) 2018 The R Foundation for Statistical Computing
Platform: x86_64-apple-darwin15.6.0 (64-bit)
R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.
Natural language support but running in an English locale
Viewer
R Documentation
gend, and plot labels
labs(..., title = waiver(), subtitle = waiver(), caption = waiver(),
tag = waiver())
xlab(label)
ylab(label)
ggtitle(label, subtitle = waiver())
Arguments
... A list of new name-value pairs. The name should be an aesthetic.
New R Markdown
Document
Presentation
Shiny
From Template
Title:
Untitled
Author:
Suzanne Thornton
Default Output Format:
HTML
Recommended format for authoring (you can switch to PDF or Word output anytime).
PDF
PDF output requires TeX (MiKTeX on Windows, MacTeX 2013+ on OS X, TeX Live 2013+ on Linux).
Word
Previewing Word documents requires an installation of MS Word (or Libre/Open Office on Linux).
OK
Cancel

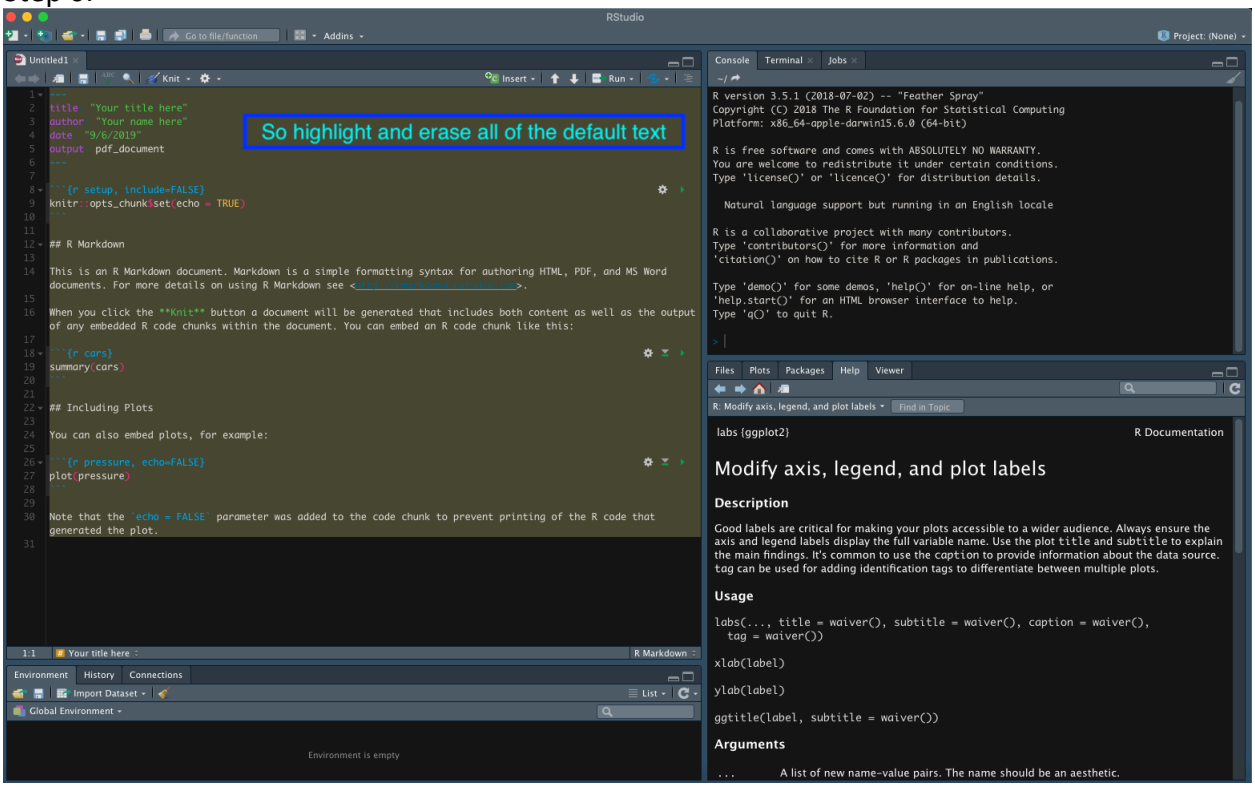
Step 4:


New R Markdown
Document
Presentation
Shiny
From Template
Title:
Your title here
Author:
Your name here
Default Output Format:
HTML
Recommended format for authoring (you can switch to PDF or Word output anytime).
PDF
PDF output requires TeX (MiKTeX on Windows, MacTeX 2013+ on OS X, TeX Live 2013+ on Linux).
Word
Previewing Word documents requires an installation of MS Word (or Libre/Open Office on Linux).
OK
Cancel

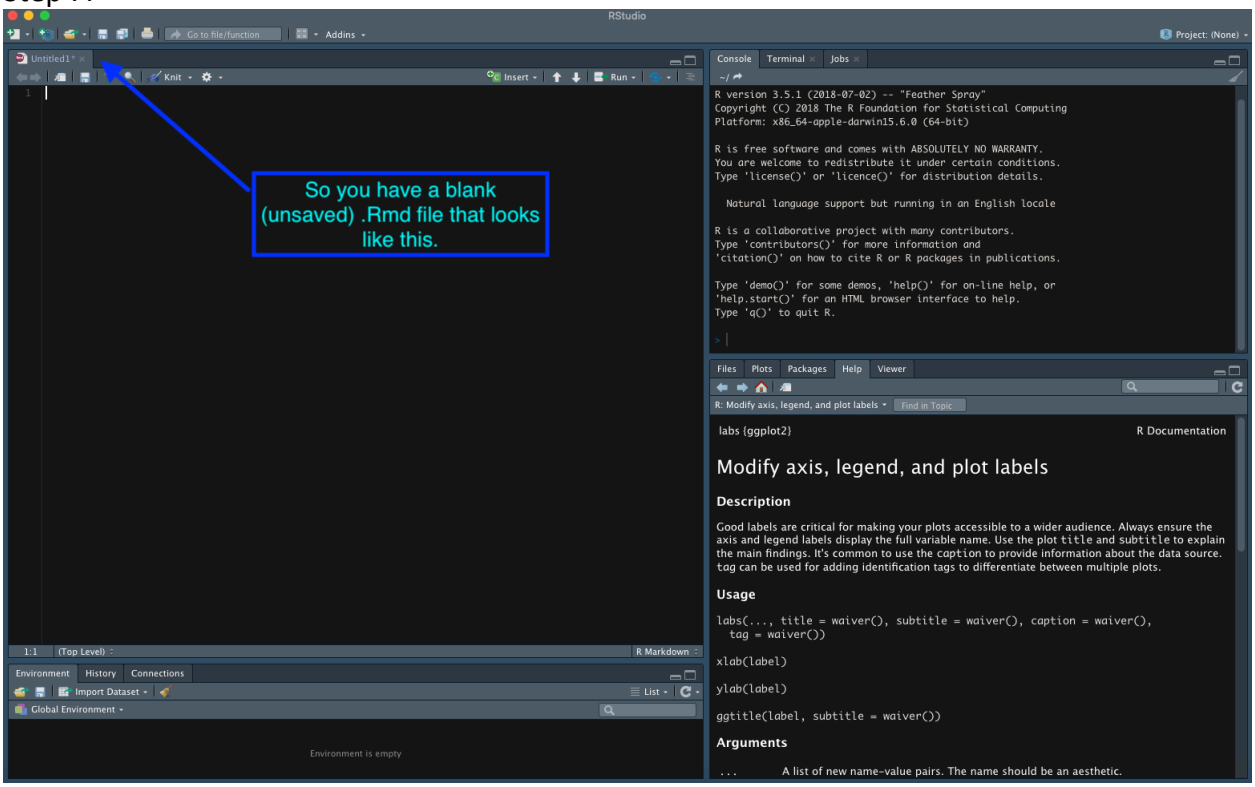
Step 5:


RStudio
R will automatically create a .Rmd template file. We won't use this however.
Modify axis, legend, and plot labels
Description
Usage
Arguments

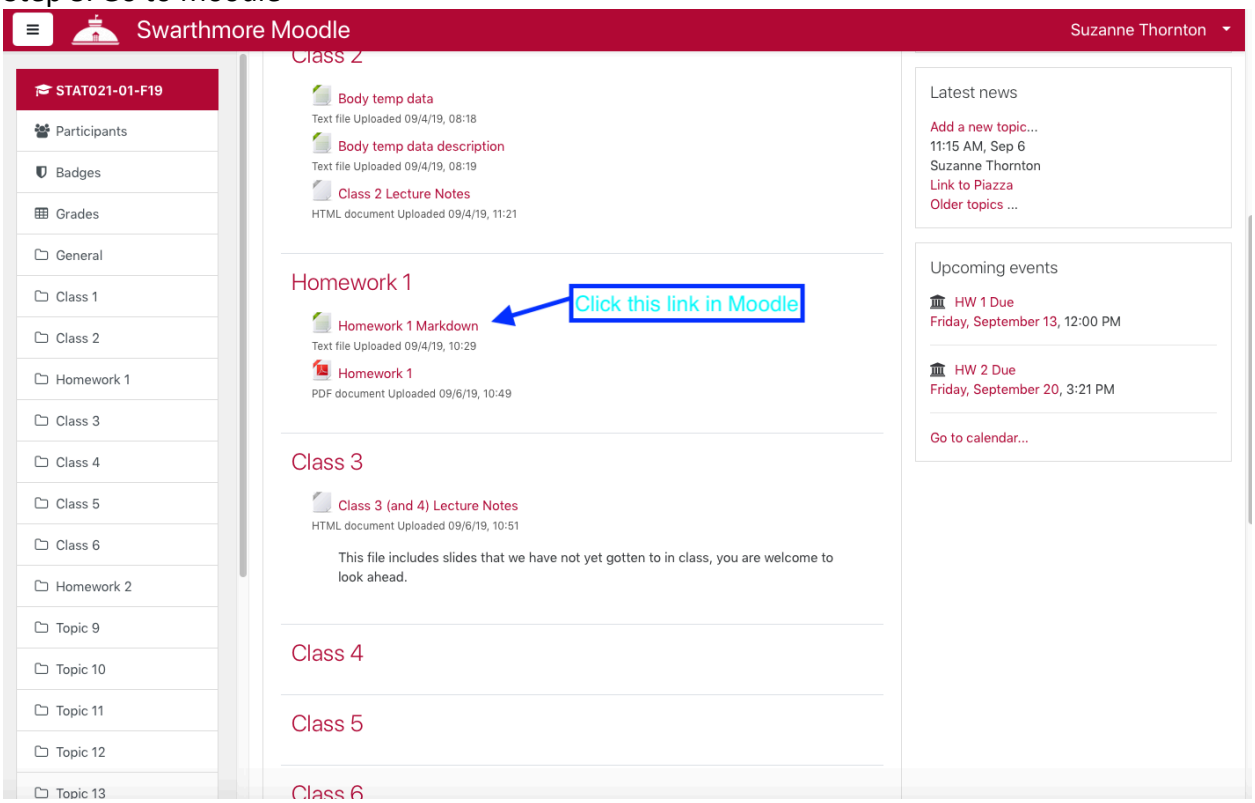
Step 6:


RStudio
So highlight and erase all of the default text
Modify axis, legend, and plot labels
Description
Usage
Arguments

Step 7:



Step 8: Go to Moodle



Step 9:

```
---
title: "Stat 021 Homework 1"
author: "Suzanne Thornton"
date: "Due: Fri, Sept 13"
output: pdf_document
---


```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
library(tidyverse)  ## includes ggplot2, dplyr, tidyr, readr, purrr, tibble, stringr, forcats
library(magrittr) ## for piping
#library(glue)  ## for easier string and variable concatination
```


<center>**Instructions:** A hard copy of your homework must be handed in to me at the end of class on the due date or I must have recieved via email a **pdf** version of your homework by **noon** on the due date. If you are having trouble getting your *.Rmd* file to compile, you need to get help with this **before** the due date.

You are allowed to hand in **only one** late homework assignment throughout the semester. If you need to hand in this particular assignment late, you must let me know via email by noon on the due date.

You are encouraged to study with your peers to help complete the homework assignments but no copying is allowed. If I see that two or more homework assignments are copied, all students involved will recieve a grade of 0 on that assignment and will forfeit (perhaps retroactively) the opportunity to hand in a late homework. </center>

***

Q 1) Download and install R and R Studio following the instructions in class. Install the package *swirl()*
using the command "install.packages("swirl")". Once the package is installed, call the package to your R session using the command "library("swirl"). Follow the instructions that pop up in your console. Select the course option "1: R Programming: The basics of programming in R" and complete the following lessons (about 20 minutes to complete each)

  - 1: Basic Building Blocks

  - 2: Workspace and Files

  - 8: Logic

Even though you don't get points for doing this, it will really help you to learn how to program in R. If the tutorial is covering something that you already know how to do, use the *skip()* command to move through the tutorial faster, but note that there are some questions which you will not be able to skip and you'll be forced to think through. (0 points)

Q 2) Design your own experimental study (made up) on a population of your choice. Provide details on what is the population under study, how will you collect a sample, what are potential biases keeping your sample from truely being random and define different treatment/control groups and describle how you will randomly assign treatments to your observational units. (5 points)

Q 3)
Access the data set called *sleep* in R. Note this data set is loaded into R automatically (you do not need to import it or install a package to access it). (10 points)

  1. Define a new variable called *group1.sleep* that includes only the values of the variable *extra* for individuals from group 1.

  1. Perform a t-test on ``group1.sleep" to test of the extra hours slept by group 1 is smaller than or equal to 0.5 hours at an $\alpha = 0.1$ level. I.e. Test the hypothesis $H_0: \mu = 0.5$ vs. $H_1: \mu \neq 0.5$ at an $\alpha = 0.1$ significance level.

  1. Report and interpret the 90% CI for the average extra hours of sleep for group 1.

  1. Form a new categorical variable called *extra1.cat* that categorizes the variable *extra* into two groups, the first where extra hours slept is at least zero hours and the
```

This will open up a new tab that looks like this

Step 10:

Highlight all of the text on this page and copy/paste it into your .Rmd file


```
---
title: "Stat 021 Homework 1"
author: "Suzanne Thornton"
date: "Due: Fri, Sept 13"
output: pdf_document
---


```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
library(tidyverse)  ## includes ggplot2, dplyr, tidyr, readr, purrr, tibble, stringr, forcats
library(magrittr) ## for piping
#library(glue)  ## for easier string and variable concatination
```


<center>**Instructions:** A hard copy of your homework must be handed in to me at the end of class on the due date or I must have recieved via email a **pdf** version of your homework by **noon** on the due date. If you are having trouble getting your *.Rmd* file to compile, you need to get help with this **before** the due date.

You are allowed to hand in **only one** late homework assignment throughout the semester. If you need to hand in this particular assignment late, you must let me know via email by noon on the due date.

You are encouraged to study with your peers to help complete the homework assignments but no copying is allowed. If I see that two or more homework assignments are copied, all students involved will recieve a grade of 0 on that assignment and will forfeit (perhaps retroactively) the opportunity to hand in a late homework. </center>

***

Q 1) Download and install R and R Studio following the instructions in class. Install the package *swirl()*
using the command "install.packages("swirl")". Once the package is installed, call the package to your R session using the command "library("swirl"). Follow the instructions that pop up in your console. Select the course option "1: R Programming: The basics of programming in R" and complete the following lessons (about 20 minutes to complete each)

  - 1: Basic Building Blocks

  - 2: Workspace and Files

  - 8: Logic

Even though you don't get points for doing this, it will really help you to learn how to program in R. If the tutorial is covering something that you already know how to do, use the *skip()* command to move through the tutorial faster, but note that there are some questions which you will not be able to skip and you'll be forced to think through. (0 points)

Q 2) Design your own experimental study (made up) on a population of your choice. Provide details on what is the population under study, how will you collect a sample, what are potential biases keeping your sample from truely being random and define different treatment/control groups and describle how you will randomly assign treatments to your observational units. (5 points)

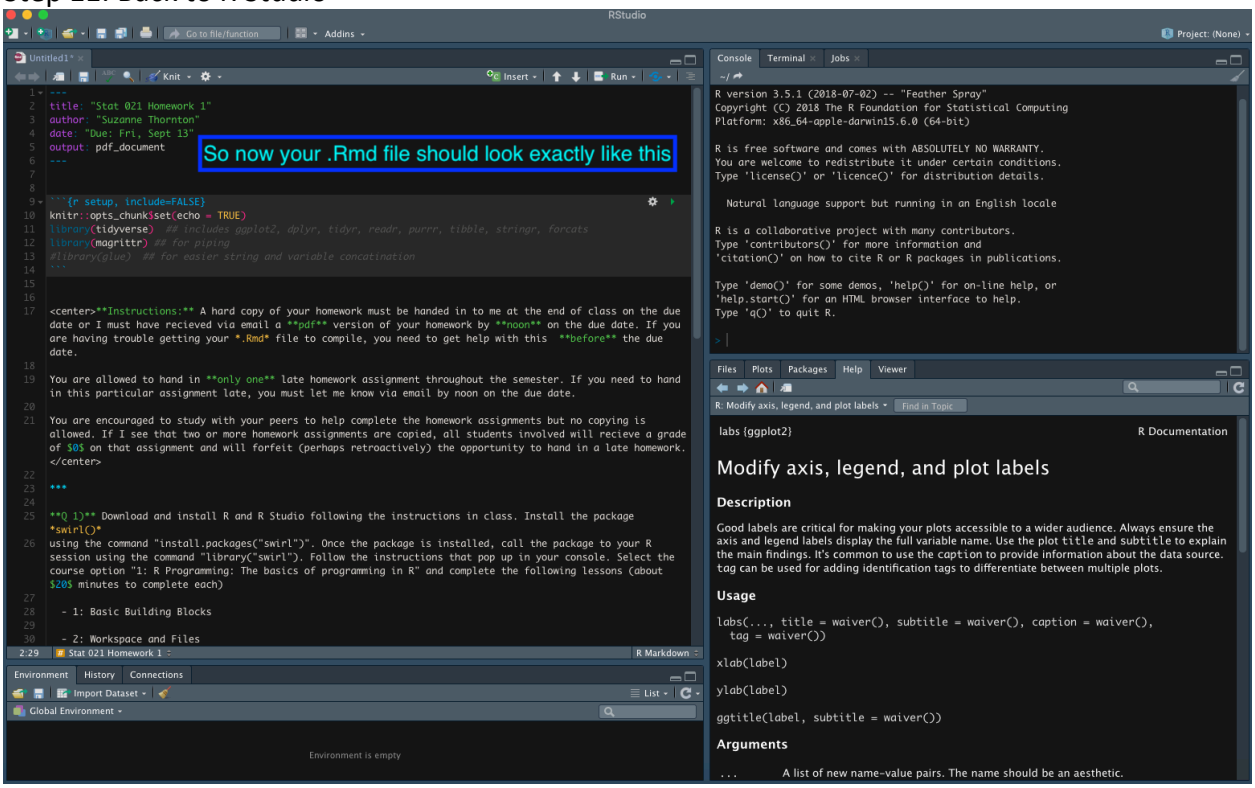
Q 3)
Access the data set called *sleep* in R. Note this data set is loaded into R automatically (you do not need to import it or install a package to access it). (10 points)

  1. Define a new variable called *group1.sleep* that includes only the values of the variable *extra* for individuals from group 1.

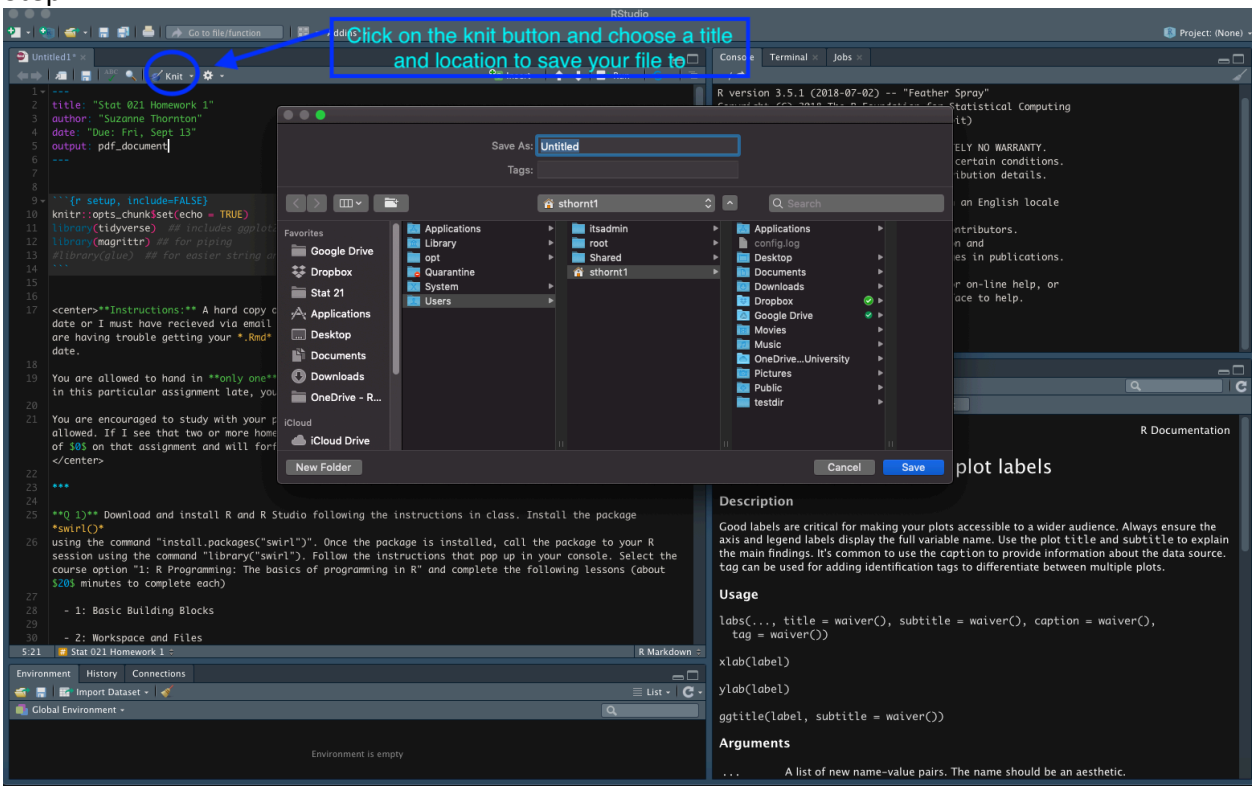
  1. Perform a t-test on ``group1.sleep" to test of the extra hours slept by group 1 is smaller than or equal to 0.5 hours at an $\alpha = 0.1$ level. I.e. Test the hypothesis $H_0: \mu = 0.5$ vs. $H_1: \mu \neq 0.5$ at an $\alpha = 0.1$ significance level.

  1. Report and interpret the 90% CI for the average extra hours of sleep for group 1.
```

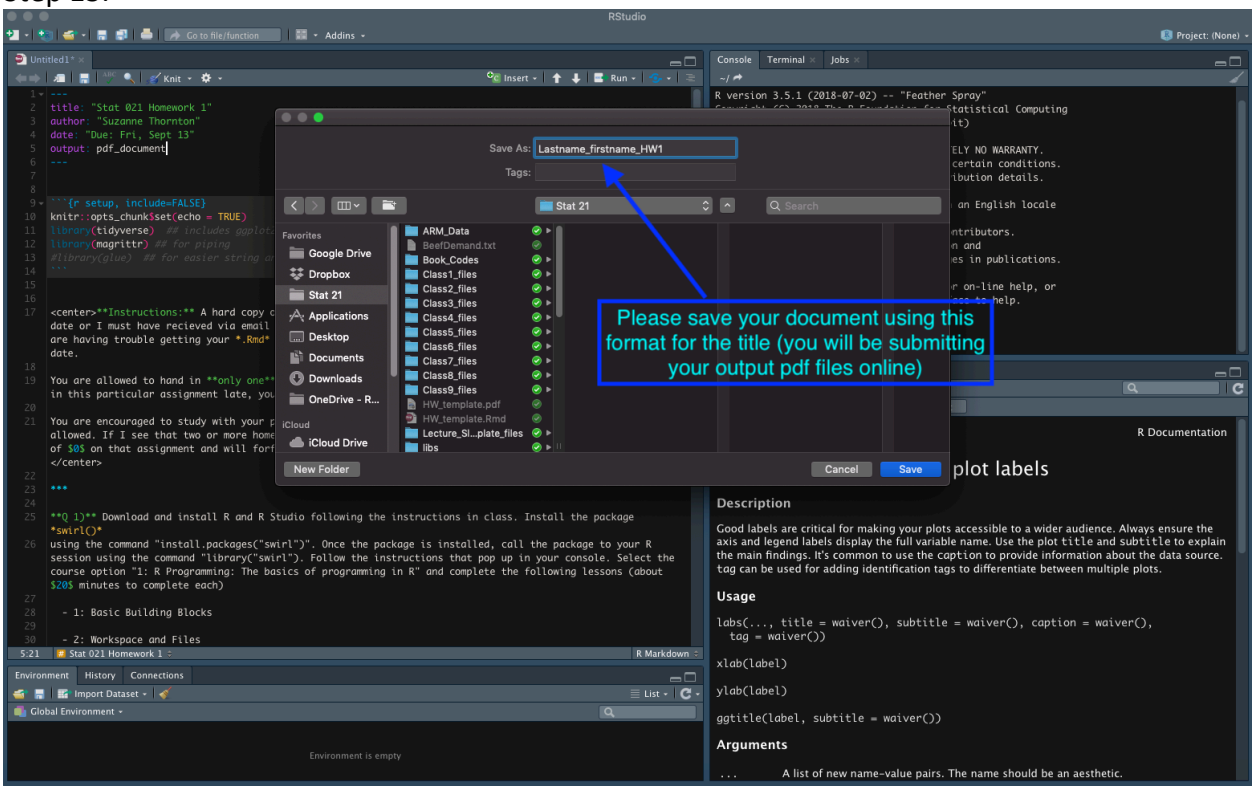
Step 11: Back to R Studio



Step 12:



Step 13:



Step 14: