

Problem Set 0: Getting Set Up!

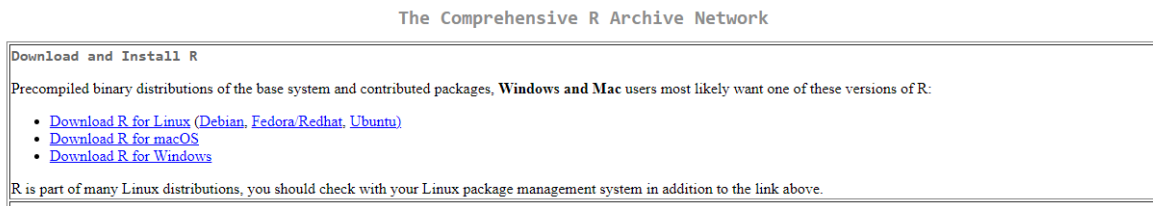
The rest of the semester will see you working on data science questions using `R`. As such, your first homework assignment is:

1. Install `R` on your computer.
2. Install `RStudio` on your computer.
3. Create a `directory` (i.e., a folder with a set of subfolders) for this class.
4. Create a new `.Rmd` file and `save` as....
5. Modify the `.Rmd` file and `knit` it.

1. Installing R

`R` is going to be the only programming language we will use. `R` is an extensible statistical programming environment that can handle all of the main tasks that we'll need to cover this semester: getting data, analyzing data and communicating data analysis.


Download R here: <https://cran.r-project.org/> (<https://cran.r-project.org/>). Make sure to choose the version that works with your operating system!



2. Installing RStudio

When we work with `R`, we communicate via the command line. To help automate this process, we can write scripts, which contain all of the commands to be executed. These scripts generate various kinds of output, like numbers on the screen, graphics or reports in common formats (pdf, word). Most programming languages have several **I**ntegrated **D**evelopment **E**nvironments (IDEs) that encompass all of these elements (scripts, command line interface, output). The primary IDE for `R` is `RStudio`.

Download `RStudio` here: <https://rstudio.com/products/rstudio/download/> (<https://rstudio.com/products/rstudio/download/>). You need the free `RStudio` desktop version.

RStudio Desktop	RStudio Desktop Pro	RStudio Server	RStudio Workbench
Open Source License	Commercial License	Open Source License	Commercial License
Free	\$995 /year	Free	\$4,975 /year (5 Named Users)
DOWNLOAD Learn more	BUY Learn more	DOWNLOAD Learn more	BUY Evaluation Learn more
Integrated Tools for R	✓	✓	✓
Priority Support	✓		✓
Access via Web Browser		✓	✓
RStudio Professional Drivers	✓		✓
Connect to RStudio Workbench  remotely	✓		
Enterprise Security			✓

3. Setting up Directories

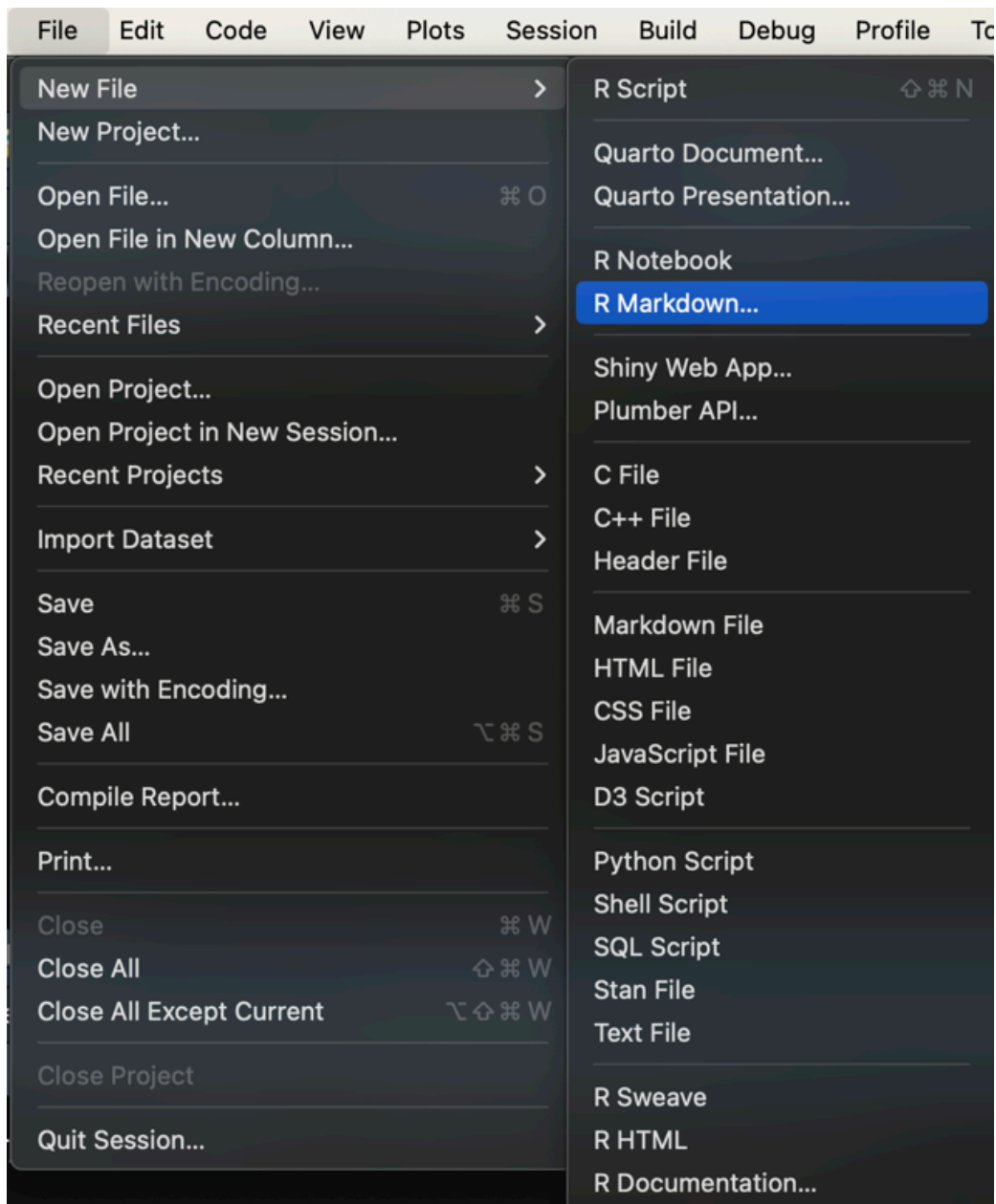
In each class, we’re going to include some code and text in one file, and data in another file. You’ll need to download both of these files to your computer. You need to have a particular place to put these files. Computers are organized using named directories (sometimes called folders). Don’t just put the files in your Downloads directory. One common solution is to create a folder on your computer named after the class: `PSC4175`.

You could just throw everything related to the class into this folder. However, this will quickly get messy. I recommend you create a “sub-folder” (or “sub-directory”) within `PSC4175` called `Lecture_1`. (You might also want to create similar sub-folders for `Lecture_2`.) Inside `Lecture_1`, create two additional sub-folders: `code` and `data`. When you’re done, your class directory should look like this:

```
- PSC4175
  - Lecture_1
    - code
    - data
```

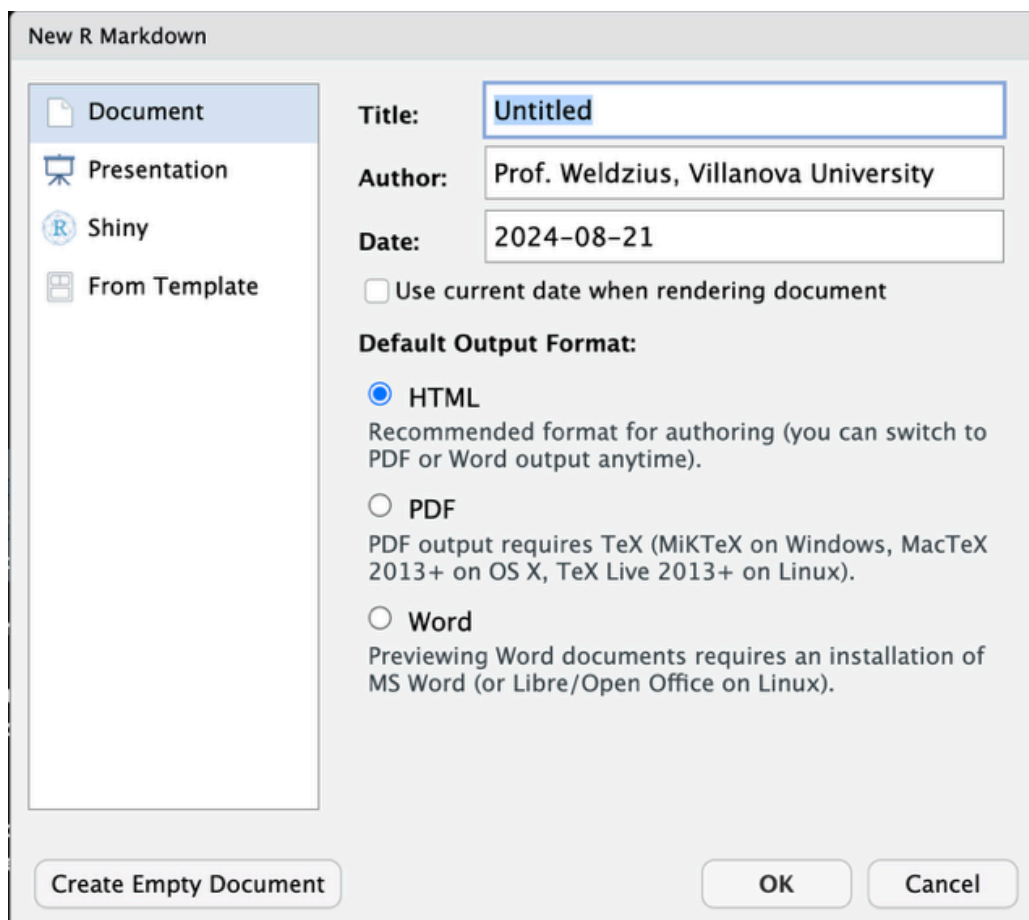
4. Create an `.Rmd` file

Open `RStudio`, then create a new `.Rmd` file. To do this, click on `File` → `New File` → `R Markdown...`



Settings for .Rmd file

You will then be asked to determine a bunch of settings for this .Rmd document. For example, you can choose whether you want to create a “Document”, “Presentation”, “Shiny”, or “From Template” on the left. You can set the “Title:” “Author:” and “Date:” on the top-right. And you can choose the “Default Output Format:” to be either “HTML”, “PDF”, or “Word”. You should **not change any of these settings**. Their defaults (“Document”, “Untitled”, “[Your name]”, “[Today’s Date]”, and “HTML”) are sufficient. Just click “OK”.



The "New R Markdown" dialog box is shown. On the left, a sidebar lists four options: "Document" (selected), "Presentation", "Shiny", and "From Template". The main area contains fields for "Title:" (Untitled), "Author:" (Prof. Weldzius, Villanova University), and "Date:" (2024-08-21). Below these is a checkbox for "Use current date when rendering document" which is unchecked. The "Default Output Format:" section has three radio buttons: "HTML" (selected), "PDF", and "Word". Descriptive text follows each format option. At the bottom are three buttons: "Create Empty Document", "OK", and "Cancel".

Title: Untitled

Author: Prof. Weldzius, Villanova University

Date: 2024-08-21

☐ Use current date when rendering document

Default Output Format:

☒ **HTML**
Recommended format for authoring (you can switch to PDF or Word output anytime).

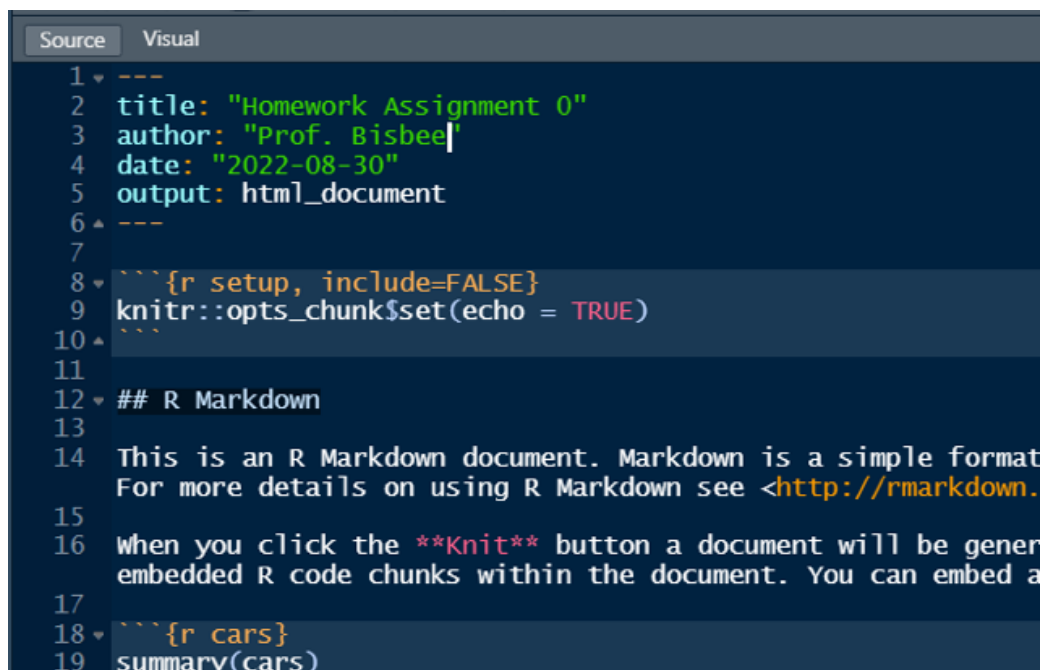
☐ **PDF**
PDF output requires TeX (MiKTeX on Windows, MacTeX 2013+ on OS X, TeX Live 2013+ on Linux).

☐ **Word**
Previewing Word documents requires an installation of MS Word (or Libre/Open Office on Linux).

Create Empty Document OK Cancel

Saving .Rmd file

This will open a new .Rmd file! Now you should change the title of the file to “Homework Assignment 0” and the author to your name. You should then save the file in your `code` folder with the file name `[Last Name]_PS0.Rmd` by clicking `File → Save As...`



The image shows a code editor window with two tabs: "Source" and "Visual". The "Source" tab is active, displaying the following R Markdown code:

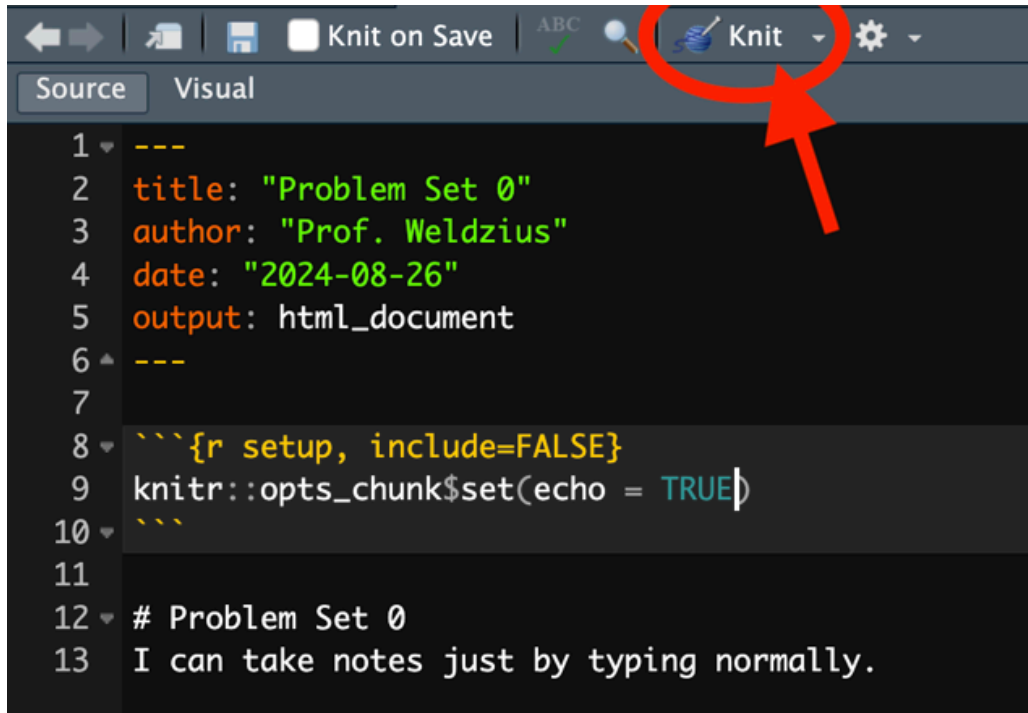
```
1 ---
2 title: "Homework Assignment 0"
3 author: "Prof. Bisbee"
4 date: "2022-08-30"
5 output: html_document
6 ---
7
8 {r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = TRUE)
10
11
12 ## R Markdown
13
14 This is an R Markdown document. Markdown is a simple format
15 For more details on using R Markdown see <http://rmarkdown.rstudio.com>
16
17 When you click the **Knit** button a document will be generated containing
18 embedded R code chunks within the document. You can embed an R code chunk
19 like this:
20
21 {r cars}
22 summary(cars)
```

5. Modify and knit

Now, **delete all of the default text** in your `.Rmd` file from line 12 down to the bottom. Then write the following on line 12:

```
# Problem Set 0
I can take notes by just typing normally.
```

Now let's `knit` the file by clicking the `knit` button on the top of the window. You should see a new window pop-up that shows the processed code! This `.html` document should also appear in your `code` folder with the same file name as your `.Rmd` file.



Inserting R Code

The final part of the homework assignment is to insert a `chunk` of `R` code. On line 15 type the following:

```
```{r}
2+2
```
```

Then `knit` a final time and open the `.html` file that is created in your `code` folder. It should be called `[Your Last Name]_PS0.html`. You should open this in your internet browser (Safari, Google Chrome, FireFox, etc.), and then right-click anywhere on the opened file and click “Print...”.

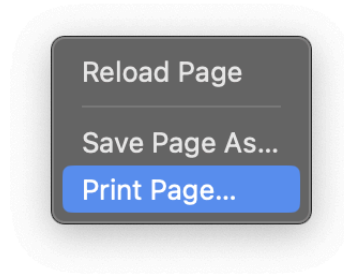
Problem Set 0

Prof. Weldzius

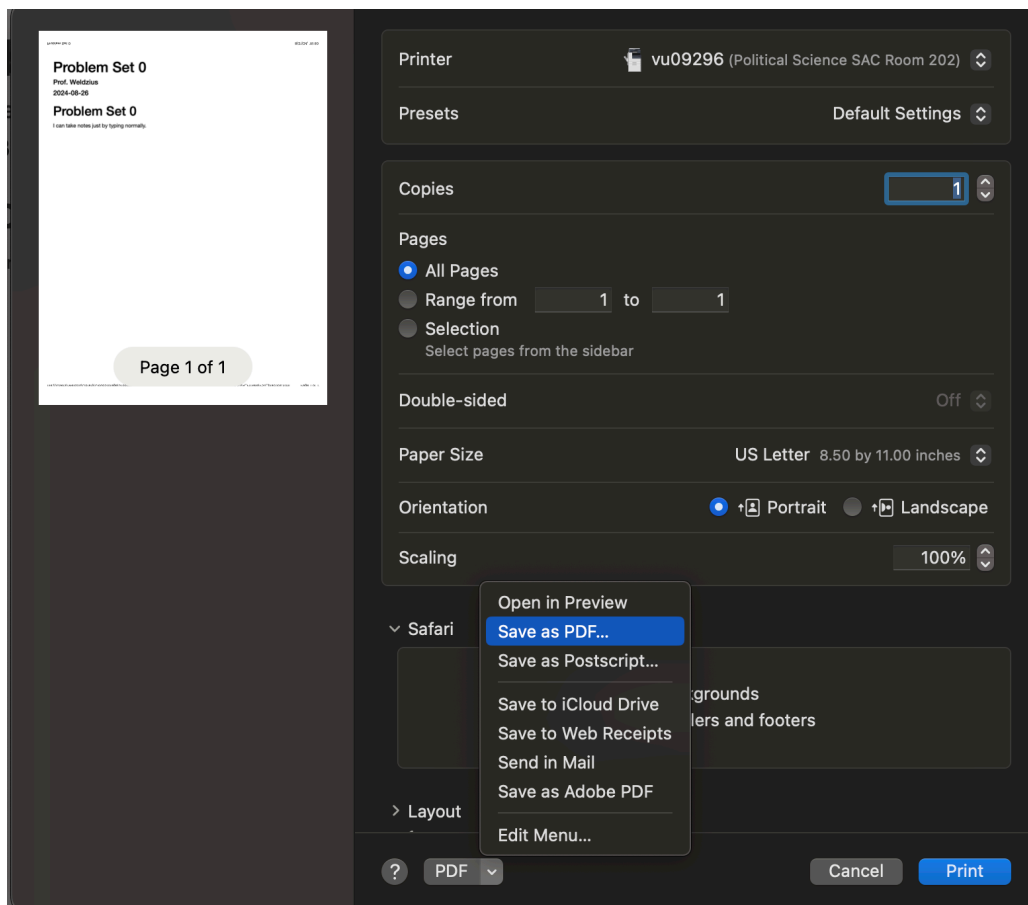
2024-08-26

Problem Set 0

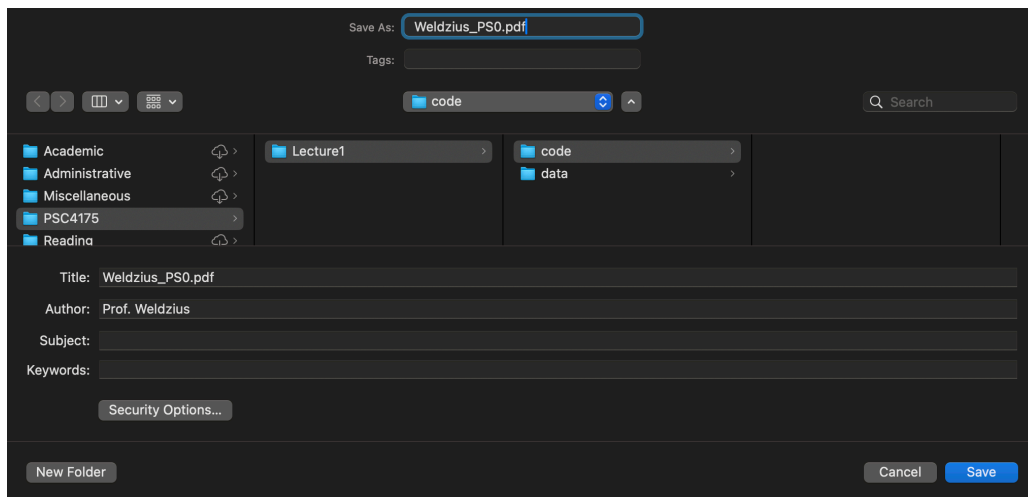
I can take notes just by typing normally.



Then choose “Save as PDF...”



and save it as [Your Last Name]_ps0.pdf .



Finally, upload the PDF to Blackboard under the Problem Set 0 assignment!