The framework for a reproducible report

497 / 597 Reproducible Research

Richard Layton

Rose-Hulman Institute of Technology
Fall 2018

Image from Wickham & Grolemund, *R for Data Science*, http://r4ds.had.co.nz/r-markdown.html

Software overview

our interface to all the software

for combining your code, its results, and your prose

for data carpentry, analysis, and data graphics

for local version control

for remote, asynchronous collaboration

Today we'll set up directories and start an Rmd script

our interface to all the software

for combining your code, its results, and your prose

for data carpentry, analysis, and data graphics

for local version control

for remote, asynchronous collaboration

To start, set up a course directory

497 students

```
me497_reproducible_research
  |-- practice_work/
  |-- project_1/
  `-- project_2/
```

597 students

```
me597_reproducible_research
  |-- practice_work/
  |-- project_1/
  |-- project_2/
  `-- project_3/
```

- the top level is the course directory
- every subfolder is a separate R project directory

  `practice_work/`

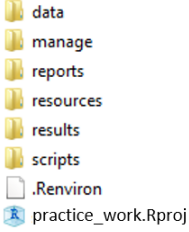
  `project_1/`

  `project_2/`

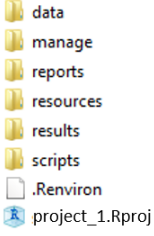
  `project_3/`

Set up your project directories consistently

practice_work/



project_1/



- create the same set of folders for every project
- an R Project directory is denoted by the `.Rproj` file
- copy the `.Renviron` file to every project

Instructions for creating an R Project and the `.Renviron` file are on the course website

To start a report

In RStudio, launch the project, for example, *practice_work.Rproj*

File > New File > R Markdown > OK

Save As to your reports directory

Be deliberate in selecting file names, for example, *001_first_script.Rmd*

The three important elements of an Rmd file

- YAML header, surrounded by ---

```
---
title: "Sample report"
author: "Richard Layton"
date: "September 1, 2016"
output: word_document
---
```

- Chunks of R code surrounded by ```

````
```{r setup, include = FALSE}
library(ggplot2)
library(dplyr)
smaller <- diamonds %>% filter(carat <= 2.5)
```
````

- Text mixed with simple text formatting like `# heading` and `*italics*`

Writing an R Markdown script

R Markdown script

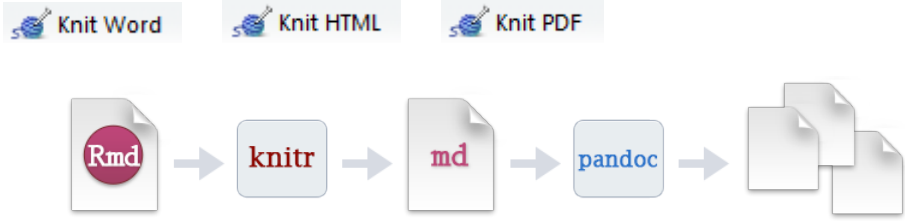
````
---
title: "Sample report"
author: "Richard Layton"
date: "September 1, 2016"
output: word_document
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

## R Markdown

This is an R Markdown document. Markdown is a simple formatting
syntax for authoring HTML, PDF, and MS Word documents. For more
details on using R Markdown see <http://rmarkdown.rstudio.com>.
````

Knit the document any time



- R Markdown sends the .Rmd file to knitr
- knitr executes the code chunks and creates a markdown (.md) document that includes the code and its output
- pandoc process the .md file to create the output file

Image from Wickham & Grolemund, *R for Data Science*, http://r4ds.had.co.nz/r-markdown.html

Sample output when knitting to Word

R Markdown script

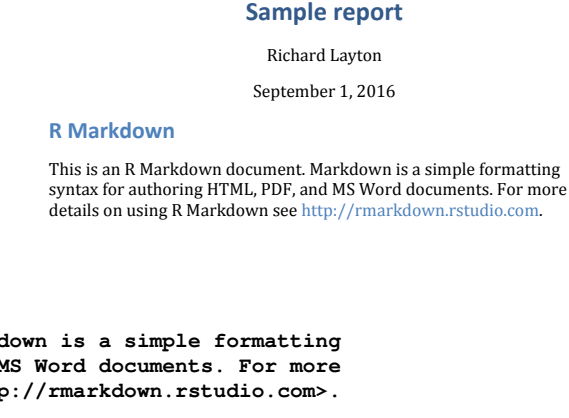
````
---
title: "Sample report"
author: "Richard Layton"
date: "September 1, 2016"
output: word_document
---

```{r setup, include=FALSE}
knitr::opts_chunk$set(echo = TRUE)
```

## R Markdown

This is an R Markdown document. Markdown is a simple formatting
syntax for authoring HTML, PDF, and MS Word documents. For more
details on using R Markdown see <http://rmarkdown.rstudio.com>.
````

Output document



Sample report

Richard Layton

September 1, 2016

R Markdown

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see http://rmarkdown.rstudio.com.

Add a code chunk

script

```
When you click the **Knit** button
a document will be generated that
includes both content as well as the
output of any embedded R code chunks
within the document. You can embed
an R code chunk like this:

```{r}
summary(cars)
```
```

Add a code chunk

script

```
When you click the **Knit** button
a document will be generated that
includes both content as well as the
output of any embedded R code chunks
within the document. You can embed
an R code chunk like this:

```{r}
summary(cars)
```
```

output

When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document. You can embed an R code chunk like this:

```r
summary(cars)
```

```
##      speed           dist
##  Min.   : 4.0   Min.   :  2.00
##  1st Qu.:12.0   1st Qu.: 26.00
##  Median :15.0   Median : 36.00
##  Mean   :15.4   Mean   : 42.98
##  3rd Qu.:19.0   3rd Qu.: 56.00
##  Max.   :25.0   Max.   :120.00
```

Add a graph

script

```
## Including Plots

You can also embed plots, for example:

```{r echo=FALSE}
plot(pressure)
```

Note that the `echo = FALSE` parameter
was added to the code chunk to prevent
printing of the R code that generated
the plot.
```

Add a graph

script

````
## Including Plots

You can also embed plots, for example:

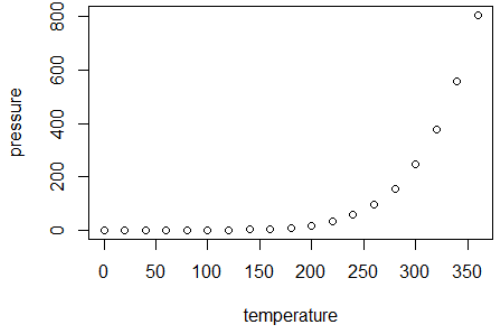
```{r echo=FALSE}
plot(pressure)
```

Note that the `echo = FALSE` parameter
was added to the code chunk to prevent
printing of the R code that generated
the plot.
````

output

Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.

Review the course website

Calendar for assignments and due dates

- projects
- practice work
- readings in the text
- articles for class discussion