Table of Contents

# Dissertating in RMarkdown + Bookdown: A preliminary guide

This tutorial was last updated: 21 April, 2019

This is a non-exhaustive guide to writing your dissertation using RMarkdown and Bookdown. Specifically, it will walk you through *one method* of organziing, writing, and rendering a dissertation with these tools, using an adapted version of the [Western University](uwo.ca) [thesis templates](https://grad.uwo.ca/academics/thesis/formatting.html). This tutorial was written by me, [Thea Knowles](theaknowles.com). At the time of writing, I am currently in the throes of dissertating. This means that there are likely several details I haven’t quite hammered out yet, or techniques I’ve missed. In the last year and a half, I’ve been collecting other people’s tutorials and resources on using RMarkdown + for the purposes of using it to write a dissertation. The final product is my interpretation of these resources, adapted to my needs, and presented here as a *“What-I’ve-learned-so-far”*-style tutorial.

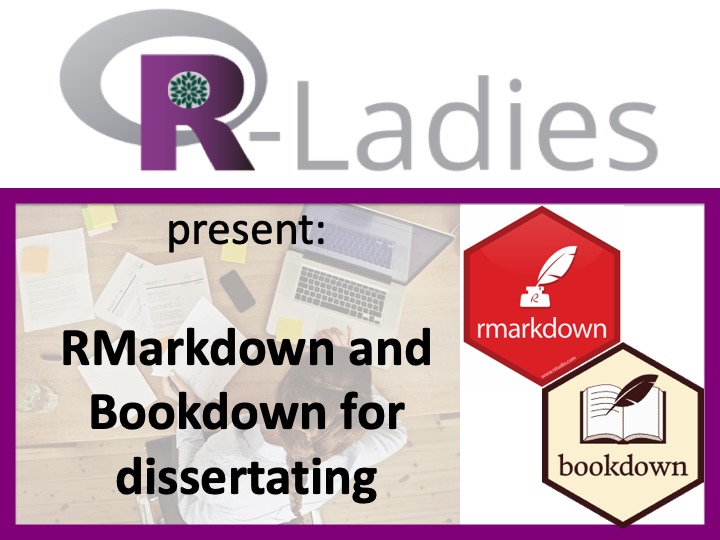
## Prerequisites

In order to use this tutorial, you need the following:

* [R](https://www.r-project.org/)
* [RStudio](https://www.rstudio.com/products/rstudio/download/)
  + Recent versions of RStudio also include Pandoc, which is required to compile documents
* Latex for [Mac](https://tug.org/mactex/mactex-download.html) or [Windows](https://miktex.org/download) (if you want to compile to PDF).
  + Alternatively, install [TinyTex](https://yihui.name/tinytex/), the Latex distribution created and recommended by Yihui Xie, creator of RMarkdown and bookdown[[1]](#footnote-31).
* R packages:

if(!require(devtools))  
 install.packages("devtools", repos = "http://cran.rstudio.com")  
install.packages("bookdown")  
install.packages("knitr")

# Introduction



## My process

1. ✏️ Keep notes in a separate bookdown project. This also helped me cut my teeth on bookdown, and made my notes wayyyyy easier to sift through than previous attempts (Google Docs, actual notebooks, txt files, readmes… I have found a lot of ways to flail)
2. 🗒 Get some data in a .csv file
3. 🧹 R file #1: load, tidy, and explore data
4. 📈 R file #2: stats and prep data for use in text, sourcing the work done in R file #1.
5. 📖 Rmd file: Write up results! Figures coded in text. Stats reported from final models defined in R file #2, with help from predefined functions and code snippets.
   * This does not have to be a complete chapter. It can be part of a chapter. How modular you want to get is really up to you.
6. 💻 Preview dissertation for myself in an .html format (compiles faster, easier to navigate, and I save it to my Chrome bookmarks for quick access)
7. 📝 Preview multiple .Rmd files in Word to send to my supervisor as .docx
   * previews/tmp\_preview.Rmd uses child files to knit together a subset of my .Rmd files to send to my supervisor
   * There is also a “preview” function with Bookdown, but I haven’t been using this (you could though!)
8. 📚 When happy, compile the whole dissertation with what I have so far
   * Ensures that I am able to notice and fix any compiling errors at the interim stages and makes me feel like I am making real progress

# RMarkdown crash course

*But first: A crash course in RMarkdown*

In Spring 2018 at R-Ladies #LdnOnt we practiced making a [manuscript using RMarkdown](https://github.com/rladies/meetup-presentations_london_ontario/tree/master/2018-03-06_rmarkdown) ([slides](http://rpubs.com/thealk/368020)). Today we will up the ante and write a whole[[2]](#footnote-38) dissertation using RMarkdown!

bookdown is a brilliant R package designed to incorporate multiple RMarkdown files into a single final product, a book. But bookdown is not just for books:

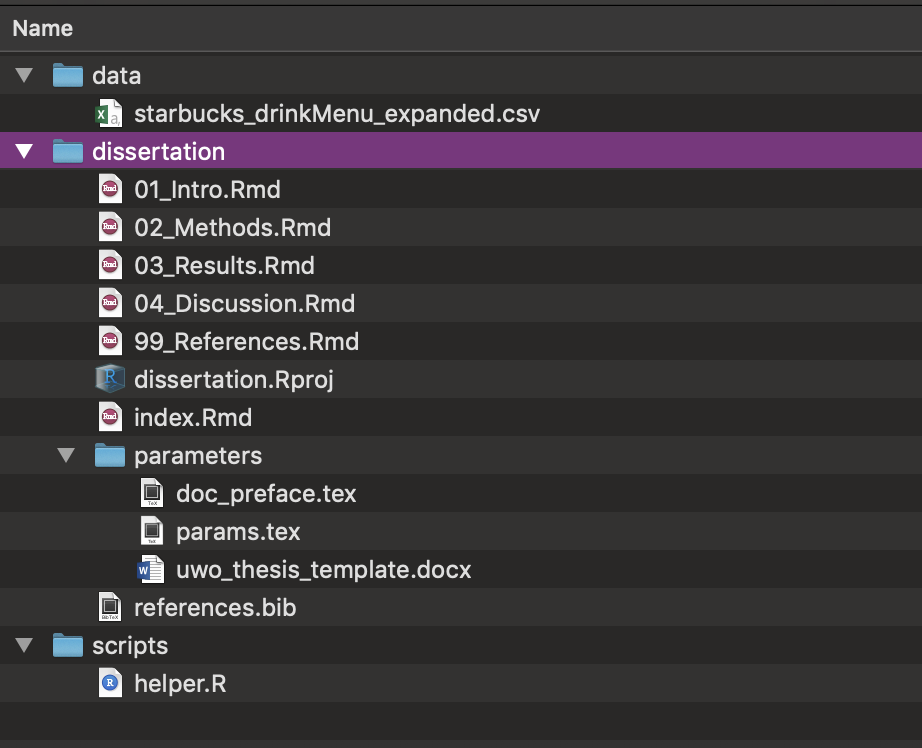
Despite the package name containing the word “book”, bookdown is not only for books. The “book” can be anything that consists of multiple R Markdown documents meant to be read in a linear sequence, such as course handouts, study notes, a software manual, a thesis, or even a diary. [From the bookdown preface](https://bookdown.org/yihui/bookdown/)

**Get up and running with RMarkdown and R Projects: Do exercise 1 in Section 5.1**

# Methods

## Getting organized: Directory structure

*A suggestion based on the advice of several clever people I have copied:*



In this example…

**dissertation** contains your individual RMarkdown files that contain the bulk of the text of the dissertation.

* index.Rmd: Contains your YAML that will tell bookdown how to render your book
* .Rmd files that comprise the body of your book. These can be specific chapters, but can also be constructed modularly in whatever way you choose.
  + Unless specifically told otherwise, bookdown will compile these in *alphanumeric order*, so they should be named in the order you want them to appear.
  + In a nutshell, file names should be **machine readable, human readable, and play well with default ordering** [(thank you, Jenny Bryan)](https://speakerdeck.com/jennybc/how-to-name-files).
  + The names of the files themselves don’t appear anywhere in the final document. For that, you need to use headers within the body of the .Rmd documents
* references.bib is a bib file containing your references. Most popular reference management tools have the option to export your references to a .bib file.
* parameters contains the files necessary to tell bookdown how to render your final document.
  + If you are compiling to a **PDF**, it needs 2 .tex (Latex) files: doc\_preface.tex, which contains the “front matter” of your dissertation (acknowledgements, etc), and params.tex, which contains the Latex parameters required to compile
  + If you are compiling to a **Word document**, it needs a template file. This could technically just be a blank document. The important thing is that it has been saved with the Word styles you want to employ in your final .docx output.

**data** contains any raw data files (.csv, .xlsx, etc.). Ideally, these don’t get touched after you put them here, because any further manipulation will be done using .R scripts (which will make it easier to track your changes to the data)

**scripts** contains any helper scripts you used along the way (e.g., for your analysis)

**More resources:**

* [Software Carpentry’s guide on project management in R Studio](https://swcarpentry.github.io/r-novice-gapminder/02-project-intro/)
* [Angela Li’s thread on thesis structuring](https://twitter.com/CivicAngela/status/1024469727274565633)

## Writing a chapter

We will now create a single chapter in RMarkdown.

There is a *lite* and a *heavy duty* version of this.

The lite version will help us learn:

* the minimal components of a bookdown chapter

**Go to the lite version in exercise 2 in Section 5.2**

The heavy-duty version will help us learn how to incorporate:

* data
* helper .R script
* citations
* figures/tables

**Go to the heavy-duty version in exercise 3 in Section 5.3**

## Making a book!

Now that we have the bare bones of a dissertation, we can compile it for the first time.

Recall that, for this method, the **essential ingredients** are:

* your .Rmd files (chapters)
* index.Rmd

**Optional ingredients:**

* references.bib
* templates

We’re going to compile in Word

### 4. Intermediary stages

* Previews (with bookdown, or just with Rmarkdown)

### 5. The nitty gritty

In descending order of messiness (i.e., how confused I get[^2])

* Reference management + citations
* Notes to self within text
* Snippets! 💕
* Predefined functions
* Footnotes
* Cross-referencing sections/figs/tables
* Figure/table autonumbering
* Tables in RMarkdown
  + Specifically when working with both .docx and .pdf
* Appendices

## Let’s start dissertating with a .docx output

### Template

* Western has a .docx thesis template. This includes detailed descriptions of what should go in the final document, but importantly for our purposes here, all we need are the *styles* specified within this document.
  + This could technically just be a blank document. We use it to tell RMarkdown[[3]](#footnote-52) how to style the final output in Word
  + See also custom\_template.docx from previous [RMarkdown workshop](LINK) to see what I mean

## Other resources

* Other Rmd + bookdown resources

## Dissertating in Bookdown

### Essential ingredients

* index.Rmd: Contains your YAML that will tell bookdown how to render your book
* .Rmd files that comprise the body of your book. These can be specific chapters, but can also be constructed modularly in whatever way you choose.
  + Unless specifically told otherwise, bookdown will compile these in *alphanumeric order*, so they should be named in the order you want them to appear.
    - In a nutshell, file names should be **machine readable, human readable, and play well with default ordering** [(thank you, Jenny Bryan)](https://speakerdeck.com/jennybc/how-to-name-files).

### Non-essential but very helpful ingredients

* A .bib file contianing your references
* Separate directories containing your:
  + data (raw .csv files, etc)
  + scripts (helper .R scripts that contain the bulk work of your analyses)
  + images that are not created in R
* **Template documents** that will provide bookdown with the information it needs to style your document
  + **Word**: uwo\_thesis\_template.docx is a Word document containing the style settings required to output a Western-approved thesis. This template also contains text, but really all you need is a Word document with the style elements set that you wish to use
  + **PDF**: The tex folder contains two .tex files that use Latex to create a final PDF:
    - params.tex contains the Latex parameters required to compile the document
    - doc\_preface.tex contains the *front matter*: everything that will appear before the beginning of the thesis

*A note on the .tex documents:* The documents provided here are comprised of a mash-up of two extremely useful resources:

* [Lucy D’Agostino McGowan’s blogpost and dissertation toolkit](https://livefreeordichotomize.com/2018/09/14/one-year-to-dissertate/) provided the original versions of these documents and a super handy walkthrough on how to use them.
* [Jon Clau’s updated Western .tex templates](https://github.com/jclauneuro/thesis_template) contain the information necessary to produce a Western U thesis (the files hosted on [Western’s site](https://grad.uwo.ca/academics/thesis/formatting.html) are older versions)

### Rendering the book

Once you have more than one .Rmd file, you are ready to render/compile them into a fully combined document.

The rendered version will appear in \_book/

To render the book as whatever is listed as the default in index.Rmd, do:

bookdown::render\_book("index.Rmd")

To render as a pdf or word document, do:

bookdown::render\_book("index.Rmd", output\_format = "pdf\_book")  
bookdown::render\_book("index.Rmd", output\_format = "word\_document2")

A nice way to preview it is to render it as a gitbook:

bookdown::render\_book("index.Rmd", "bookdown::gitbook")

Then open index.html in your browser, and run bookdown::serve\_site() to update it whenever you save changes. This is much neater/navigable.

Beware of features that will only show up in pdf rendering (like certain features of kable, kableExtra), or that won’t compile for pdfs (like the emo package)

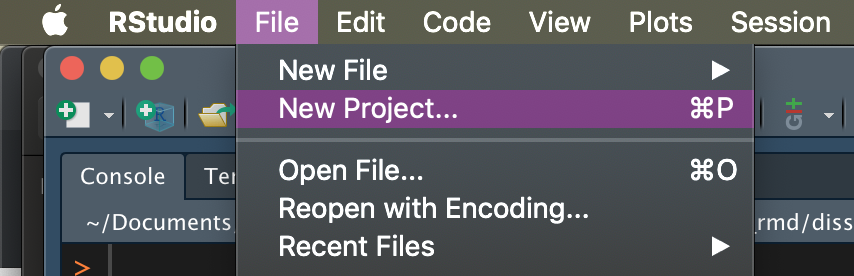
# Exercises

## 1. Getting started with RMarkdown

**Make an R Project**

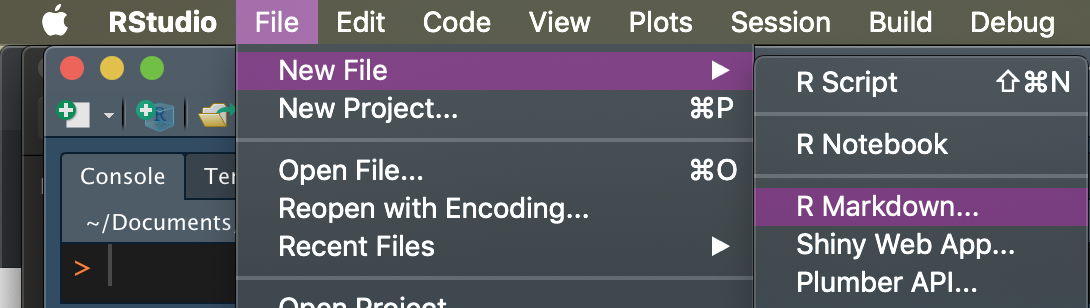
[R Projects](https://support.rstudio.com/hc/en-us/articles/200526207-Using-Projects) make project management really simple. For every new project you embark on, creating a new .RProj file. Open that .RProj file whenver you’re ready to work on that project, and it will:

* Give you easy access to the directory structure (no need to define a working directory)
* Restore your last RStudio session from that project (no need to reopen files)
* Give you access to your R history from your last session
* *NB: these preferences can all be tweaked*

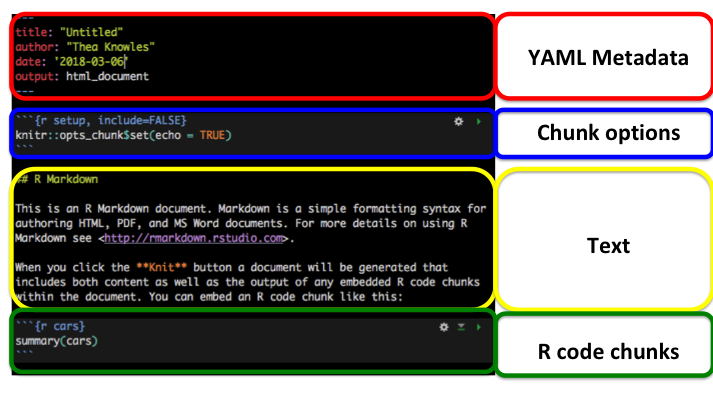


* Follow the prompts to create a project in a new directory for this workshop

**Make a new .Rmd file**

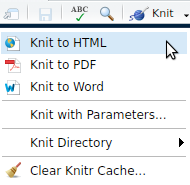


These are the components of your RMarkdown file:



We won’t go into this in *too* much detail

**Knit your document**



Try it as a…

* Word document
* Html document
* PDF document

**More resources**

* [RStudio intro to RMarkdown](https://rmarkdown.rstudio.com/lesson-1.html)
* R-Ladies #LdnOnt presentations
  + [Intro to RMarkdown](https://github.com/rladies/meetup-presentations_london_ontario/tree/master/2017-03-04_Intro2RMarkdown)
  + [RMarkdown for summary reports and journal articles](https://github.com/rladies/meetup-presentations_london_ontario/tree/master/2018-03-06_rmarkdown) (also see more specific resources on the last slide)

Back to Section 3

## 2. What’s in a chapter? Lite version

Open 01\_Intro.Rmd and do the following:

* Write some text
* Insert an R chunk that contains some simple R code
* Knit to HTML

E.g.,

x = 3  
x^2

## [1] 9

## 3. What’s in a chapter? Heavy-duty version

Open 03a\_Results.Rmd and find examples of:

* sourcing a helper file (helper.R)
* citing references
* generating images and tables from the imported data
* cross-referencing those figures and tables
* using inline R code to refer to values from the data (like p-values, etc) so there’s never a need to copy/paste

Open ../scripts/helper.R to see what it contains.

Open 03b\_Results.Rmd and do any of the following:

* source helper.R
* write some text
* cite another source
* make a figure or a table
* Knit the document as an HTML file

**More resources:** - [Getting started with bookdown](https://bookdown.org/yihui/bookdown/get-started.html)

## 4. Explore .tex files

## 5. Bookdown to .pdf

## 7. Add figure

## 9. Add table

## 10. Edit refs

## 11. Report stats

## 12. Customize snippets

# Results

# Limitations

*I.e., the things I know that I do not yet know*

## Other methods

As mentioned, this is one particular way of doing things. You may also wish to check out R packages people have created for their theses

* [thesisdown](https://github.com/ismayc/thesisdown): Thesis template designed for Reed College.
  + [Gitbook output](https://thesisdown.netlify.com/)
  + [PDF output](https://github.com/ismayc/thesisdown_book/blob/gh-pages/thesis.pdf)
* Many others have developed customized versions of this for their universities. See the thesisdown site for a list of other available templates.
* Perhaps the Western templates could eventually be incorporated in this

## Woe is Word…

Many of the issues I run into have to do with strange behaviour in Microsoft Word. For example:

### Table formatting

Table formatting is ugly with kable(). There are alternatives to kable that are specifically designed for .docx output. These are lovely alternatives to use if you are only outputting to Word (or HTML), but I have repeatedly run into problems with these other options playing nice when alternating between .pdf and .docx outputs.

Table formatting packages for .docx outputs:

* [flextable](https://cran.r-project.org/web/packages/flextable/vignettes/overview.html)
* [captioner](https://cran.r-project.org/web/packages/captioner/vignettes/using_captioner.html)
* [kableExtra::as\_image()](https://haozhu233.github.io/kableExtra/save_kable_and_as_image.html)

# References

1. TinyTex is probably the best way to go because Yihui always anticipates the problems we will run into, but I personally have not used it. [↑](#footnote-ref-31)
2. More like a smidgin than a whole, really. [↑](#footnote-ref-38)
3. Well, really to tell RMarkdown what to tell Pandoc… [↑](#footnote-ref-52)