

DISSERTATING IN RMARKDOWN
(Thesis format: Monograph)

by

Thea Lucille Knowles

Graduate Program in Department of Health & Rehabilitation Sciences

A thesis submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

© Thea Lucille Knowles 2019

Abstract

Keywords: keyword 1, keyword 2, keyword 3.

Planning to write is not writing. Outlining, researching, talking to people about what you're doing, none of that is writing. Writing is writing.

E. L. DOCTOROW
SOME BOOK

Acknowledgments

I'd like to acknowledge my dog, Rocky, whose enthusiasm for cheese is unrivaled. If this dissertation is even half as good as Rocky thinks cheese is, I will deem this body of work a massive success.

Contents

Abstract	ii
Acknowledgments	iv
List of Figures	vi
List of Algorithms	vii
List of Tables	viii
List of Appendices	ix
List of Abbreviations, Symbols, and Nomenclature	x
Preface	xi
1 Dissertating in RMarkdown + Bookdown: A preliminary guide	1
1.1 Prerequisites	1
2 Introduction	3
2.1 My process	3
3 RMarkdown crash course	5
4 Methods	6
4.1 Getting organized: Directory structure	6
4.2 Writing a chapter	7
4.3 Making a book!	8
4.3.1 The workhorses	8
4.3.2 4. Intermediary stages	8
4.3.3 5. The nitty gritty	8
4.4 Let's start dissertating with a .docx output	9
4.4.1 Template	9
4.5 Other resources	9
4.6 Dissertating in Bookdown	9
4.6.1 Essential ingredients	9
4.6.2 Non-essential but very helpful ingredients	10

4.6.3	Rendering the book	10
5	Exercises	12
5.1	1. Getting started with RMarkdown	12
5.2	2. What's in a chapter? Lite version	14
5.3	3. What's in a chapter? Heavy-duty version	14
5.4	4. Explore .tex files	15
5.5	5. Bookdown to .pdf	15
5.6	7. Add figure	15
5.7	9. Add table	15
5.8	10. Edit refs	15
5.9	11. Report stats	15
5.10	12. Customize snippets	15
6	Results	16
7	Limitations	17
7.1	Other methods	17
7.2	Woe is Word...	17
7.2.1	Table formatting	17
8	References	19

List of Figures

List of Algorithms

List of Tables

List of Appendices

List of Abbreviations, Symbols, and Nomenclature

You can define nomenclature from the text by using `printnomenclature`:

You can also manually specify them:

Terminology

CNR Contrast-to-noise ratio

SNR Signal-to-noise ratio

Terms Introduced by this Thesis

ABC ABC

DEF DEF

Preface

Chapter 1

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 organizing, writing, and rendering a dissertation with these tools, using an adapted version of the Western University thesis templates. This tutorial was written by me, Thea Knowles. 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.

1.1 Prerequisites

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

- R
- RStudio
 - Recent versions of RStudio also include Pandoc, which is required to compile documents
- Latex for Mac or Windows (if you want to compile to PDF).
 - Alternatively, install TinyTex, the Latex distribution created and recommended by Yihui Xie, creator of RMarkdown and bookdown¹.
- R packages:

¹TinyTex is probably the best way to go because Yihui always anticipates the problems we will run into, but I personally have not used it.

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

Chapter 2

Introduction



2.1 My process

0. Keep notes in a separate bookdown project. This also helped me cut my teeth on bookdown, and made my notes wayyyyyy easier to sift through than previous

attempts (Google Docs, actual notebooks, txt files, readmes... I have found a lot of ways to flail)

1. Get some data in a .csv file
2. R file #1: load, tidy, and explore data
3. R file #2: stats and prep data for use in text, sourcing the work done in R file #1.
4. 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.
5. Preview dissertation for myself in an .html format (compiles faster, easier to navigate, and I save it to my Chrome bookmarks for quick access)
6. 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!)
7. 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

Chapter 3

RMarkdown crash course

But first: A crash course in RMarkdown

In Spring 2018 at R-Ladies #LdnOnt we practiced making a manuscript using RMarkdown (slides). Today we will up the ante and write a whole¹ 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

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

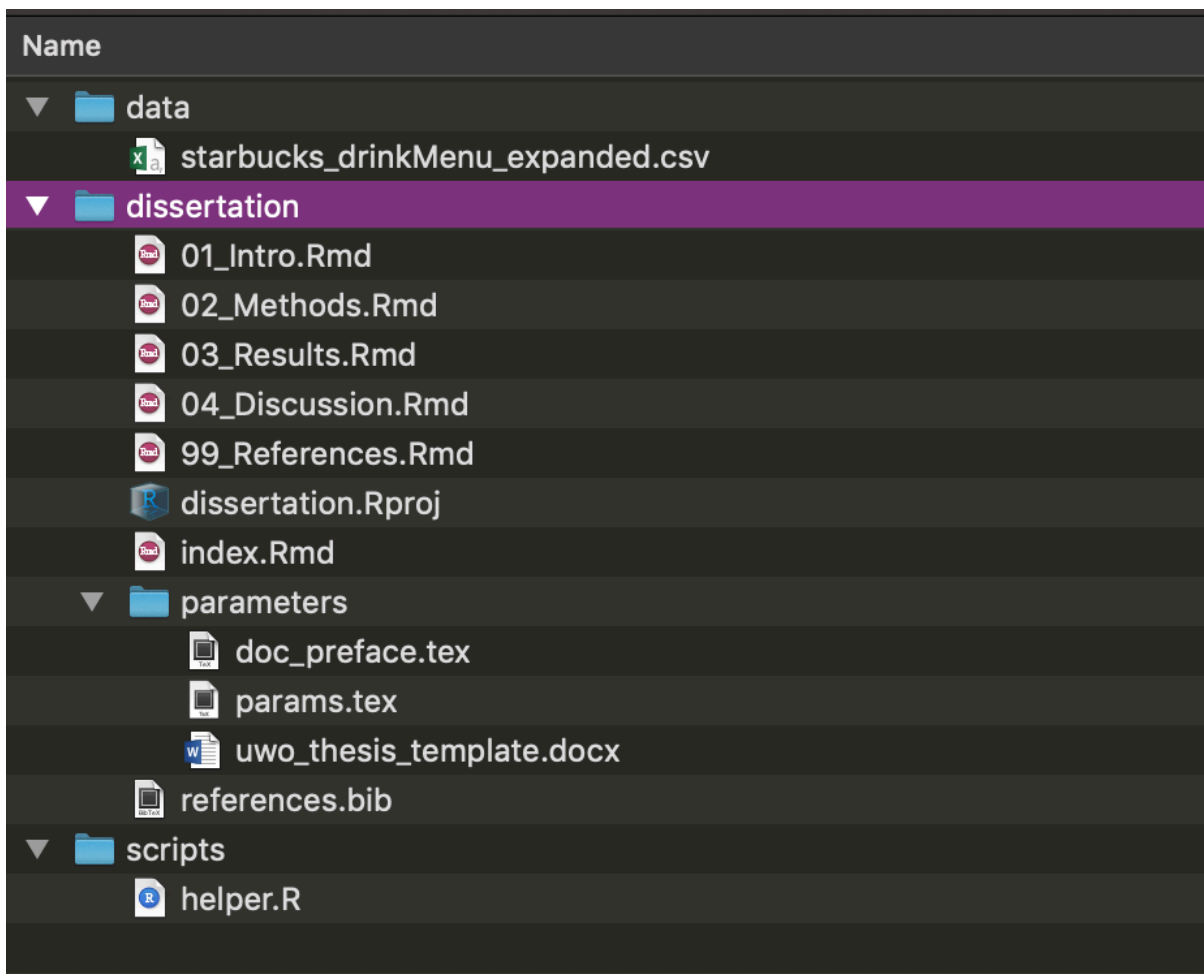
¹More like a smidgin than a whole, really.

Chapter 4

Methods

4.1 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).
 - 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
- Angela Li’s thread on thesis structuring

4.2 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

4.3 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

4.3.1 The workhorses

The files that will do the heavy lifting in order for your dissertation to compile are:

- `index.Rmd` contains the YAML¹

4.3.2 4. Intermediary stages

- Previews (with `bookdown`, or just with `Rmarkdown`)

4.3.3 5. The nitty gritty

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

- Reference management + citations
- Notes to self within text

¹YAML (rhymes with camel): The header that tells R Markdown how to generate your document. Indentation and spacing are very important. Stands for

- Snippets!
- Predefined functions
- Footnotes
- Cross-referencing sections/figs/tables
- Figure/table autonumbering
- Tables in RMarkdown
 - Specifically when working with both .docx and .pdf
- Appendices

4.4 Let's start dissertating with a .docx output

4.4.1 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 RMarkdown² how to style the final output in Word
 - See also `custom_template.docx` from previous RMarkdown workshop to see what I mean

4.5 Other resources

- Other Rmd + bookdown resources

4.6 Dissertating in Bookdown

4.6.1 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 *alphabetic 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).

²Well, really to tell RMarkdown what to tell Pandoc...

4.6.2 Non-essential but very helpful ingredients

- A `.bib` file containing 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 provided the original versions of these documents and a super handy walkthrough on how to use them.
- Jon Clau’s updated Western `.tex` templates contain the information necessary to produce a Western U thesis (the files hosted on Western’s site are older versions)

4.6.3 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)

Chapter 5

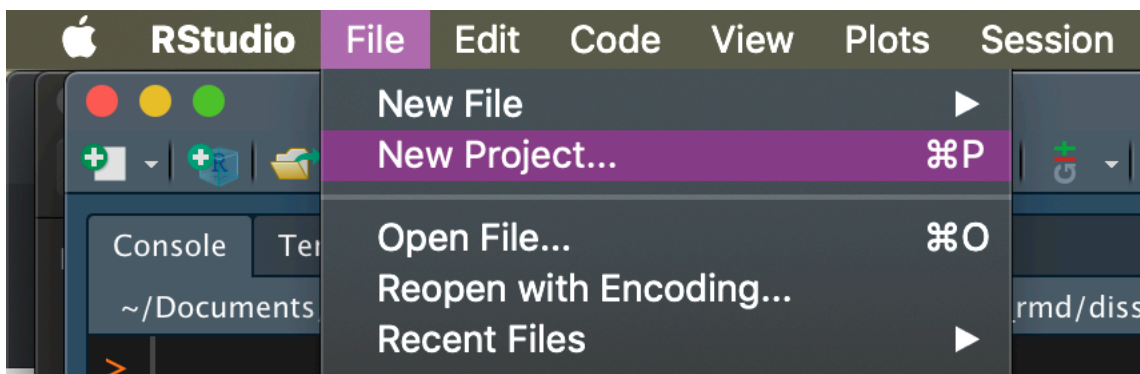
Exercises

5.1 1. Getting started with RMarkdown

Make an R Project

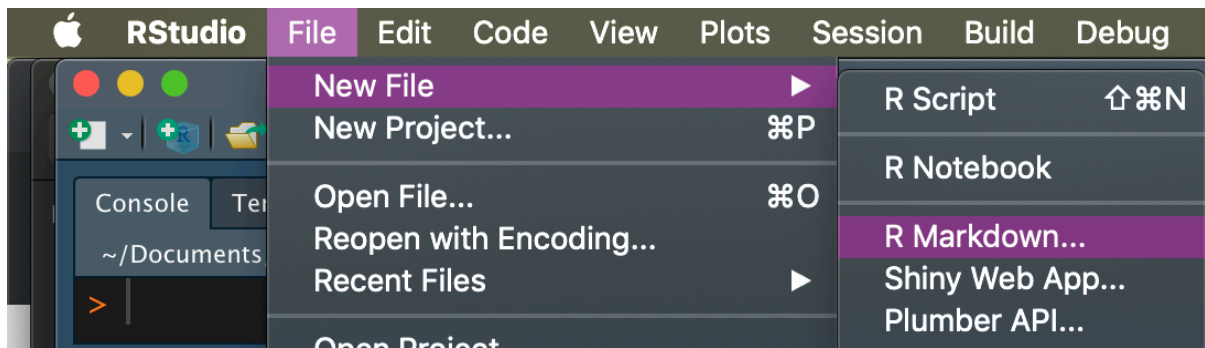
R Projects make project management really simple. For every new project you embark on, creating a new .RProj file. Open that .RProj file whenever you're ready to work on that project, and it will:

- 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*

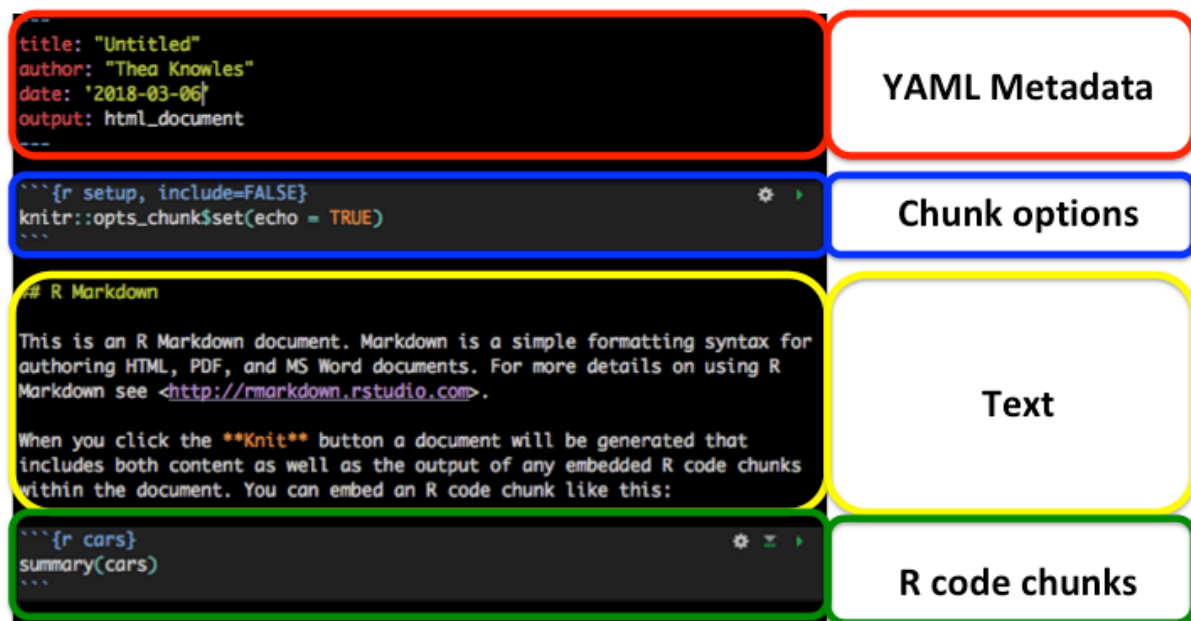


- Follow the prompts to create a project in a new directory for this workshop

Make a new .Rmd file

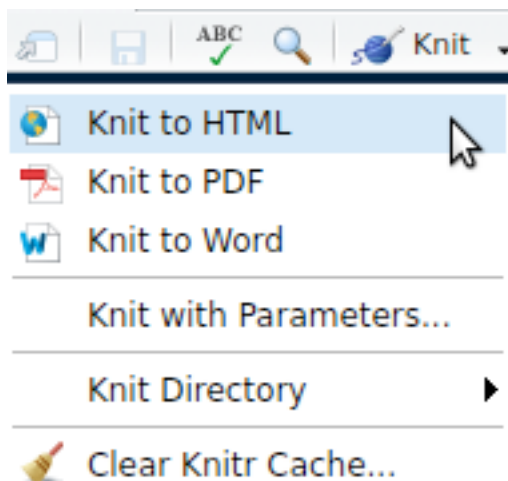


These are the components of your RMarkdown file:



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
- R-Ladies #LdnOnt presentations
 - Intro to RMarkdown
 - RMarkdown for summary reports and journal articles (also see more specific resources on the last slide)

Back to Section 3

5.2 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
```

5.3 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

5.4 4. Explore .tex files

5.5 5. Bookdown to .pdf

5.6 7. Add figure

5.7 9. Add table

5.8 10. Edit refs

5.9 11. Report stats

5.10 12. Customize snippets

Chapter 6

Results

Chapter 7

Limitations

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

7.1 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`: Thesis template designed for Reed College.
 - Gitbook output
 - PDF output
- 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

7.2 Woe is Word...

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

7.2.1 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
- captioner
- kableExtra::as_image()

Chapter 8

References