

1 Aims of the Chapter
2 What is R Markdown
3 Useful Resources
4 Using R Markdown
5 R Markdown Output type
6 Creating R Markdown files
7 Using the <code>bookdown</code> package
8 Knitting R Markdown/Notebook files
9 R Markdown syntax
10 Alternative output formats
10.1 pdf documents
10.2 Word documents
11 Exercises A
References

Chapter 6 - R Markdown

1 Aims of the Chapter

By the end of this chapter you should know

- What R Markdown is
- What R Markdown can do
- How R Markdown can make your work transparent and easy to share with others
- How to produce documents using R Markdown

Note that, for this module, we **insist** that you use R Markdown (or knitr if you prefer) to produce reports for assignments that involve writing R code. So you need to be able to use it confidently. The good news is that R Markdown is now very easy to use. All the html files used in the R programming part of this module were produced in R Markdown.

Xie, Allaire, and Golemund (2018) is *the* manual for R Markdown. What's more it has been made freely available here (<https://bookdown.org/yihui/rmarkdown/>)

2 What is R Markdown

R Markdown allows you to incorporate R code into a whole range of documents, web pages, presentations etc. Some advantages are:

- it allows you to write **simple** text that is then translated in a range of formats (html, pdf, slide etc) so you don't need to know html to produce html files
- the same R markdown file can be used to produce all the different file types (with a simple modification to the file header)

- it makes your work highly reproducible; you only have to share a single document and the reader can reproduce your output exactly
- you only have one document to work with; rather than having to run code in R, export figures from R and then read them into (a single click in R Markdown does all this)
- you can easily incorporate into R Markdown documents

To see the sort of thing that R Markdown can do have a look at this web page (<https://rmarkdown.rstudio.com/gallery.html>). We are mostly going to be using R Markdown to produce html documents. Pdfs can also be produced but can be more difficult initially (especially on University computers).

3 Useful Resources

There are excellent resources available here (<https://rmarkdown.rstudio.com/lesson-1.html>) There is a Cheat Sheet available that gives a recap of the of commonly used options in R Markdown. It is available here (<https://www.rstudio.com/wp-content/uploads/2016/03/rmarkdown-cheatsheet-2.0.pdf>) There are also more details available here (<https://www.rstudio.com/wp-content/uploads/2015/03/rmarkdown-reference.pdf>)

4 Using R Markdown

We will look at basic R Markdown which should be enough for the output you are likely to create in this module. To get started do the following:

- install the `rmarkdown` package if you haven't already;
- load the package into the R search path;
- check it is in your search path use `search()`

5 R Markdown Output type

You can choose what type of format you would like the document to be in. I am using HTML files this module. You state what output type you want in the header at the beginning of the document (more later).

6 Creating R Markdown files

R Markdown files all have a `.Rmd` extension. If you create them in RStudio then they will automatically have this extension - there is no need to include it in the file name. To create a new R Markdown file:

1. go to `file` → `New File` → `R Markdown`
2. the created file already contains some example syntax to help you

7 Using the `bookdown` package

If you look at the yaml header in your new notebook file you will see that the output type is `html_notebook`. This is fine for most purposes but better control and functionality can be obtained from using the `bookdown` package. So

- install the `bookdown` package now;
- in the yaml header replace `html_notebook` with `bookdown::html_document2`

8 Knitting R Markdown/Notebook files

Knitting is the term used for *compiling* the code in R Markdown. To knit the file use the 'Knit' button (ball of wool symbol) and select `Knit to HTML`.

You will be asked to save the file so choose a meaningful file title and save the file somewhere sensible (don't include the .Rmd extension, this is done automatically).

The file you have created already has some example code in so knit it now by clicking the `Knit to HTML` button.

9 R Markdown syntax

All you need to know now is what the R Markdown syntax looks like. Luckily R Markdown is very simple. Perhaps the easiest way to learn it is compare the contents of a .Rmd Markdown document with the HTML file produced.

10 Alternative output formats

10.1 pdf documents

We have been creating HTML files but it is easy enough to create pdfs too. First you need to replace `bookdown::html_document2` with `bookdown::pdf_document2` in the yaml header and remove the other parts of the yaml header related to HTML formats. Add the extra argument `keep_tex: true`. Your yaml header will look something like:

```
---
title: "Chapter 4 - R Markdown"
output:
  bookdown::pdf_document2:
    keep_tex: true
bibliography: refdatabase.bib
---
```

When you first knit this file you may get an error message relating to a log file. This probably means that there are some missing style files or other files. To overcome this problem, do the following:

- Because we have added the extra argument `keep_tex: true` in the yaml header, the .tex file that has been created will be saved.
- Open this .tex file in whatever TeX front-end you use (TeXWorks, WinEDT etc) and compile the .tex file (pdfLaTeX etc).
- You will then be asked to install various files which you should agree to do.
- You should now be able to knit the R Markdown code into a pdf from RStudio.

Note that you should only need to do this procedure once on any given computer. Once the required files have been loaded into MikTeX they won't need loading again.

10.2 Word documents

In the yaml header just replace `bookdown::html_document2` with `word_document` and knit it.

11 Exercises A

1. Knit the `R_Markdown_Syntax.Rmd` file into a pdf (this may not work easily on these PCs!)
2. Use R Markdown from now on to contain the R code for your answers to the questions in Chapter 5.
3. If you have finished all the questions then try this question (parts of this might be difficult for Data Analytics students in particular). Use R Markdown

Consider the random variables

- $Y \sim Ga(\alpha, \beta)$
- $X|Y \sim exp(Y)$

where α is the shape parameter and β is the rate parameter; see this Wiki page (https://en.wikipedia.org/wiki/Gamma_distribution) for the pdf of a random variable with a gamma probability distribution.

- Suppose we want to calculate $Pr(X > 2)$. We can do this by Monte Carlo integration (stats MSc and MMath students taking Computational Inference will see more of this next semester). MSc Data Analytics students should not worry about the details of this (but maths/stats students should). The trick is to write the probability as an expectation and replace the expectation with a sample mean.

$$Pr(X > 2) = \mathbb{E}(Z) \quad \text{where} \quad Z = 1 \quad \text{if} \quad X > 2 \quad \text{and} \quad Z = 0 \quad \text{otherwise}$$

$$\approx \frac{1}{n} \sum_{i=1}^n z_i$$

where z_1, z_2, \dots, z_n are n iid realisations of Z .

So we need to sample lots of z_i realisations. To do this we first sample a realisation of Y , then, given this value of Y , we sample a realisation of X from the conditional distribution $X|Y$. If this realisation of X is more than 2 then $z_i = 1$, otherwise $z_i = 0$. We do this multiple times and record the proportion of z_i values that are 1.

- Write some code to implement this process for $\alpha = 2$ and $\beta = 3$.
- By performing the integration directly it is possible to show that if $\alpha \in \mathbb{Z}^+$ then

$$Pr(X > 2) = \left(\frac{\beta}{2 + \beta} \right)^\alpha$$

Check your Monte Carlo estimate against the true value.

- If you're feeling ambitious verify, for $\alpha > 0$ and $\beta > 0$, the following result:

$$Pr(X > 2) = \frac{\Gamma(\alpha + 1)}{\alpha \Gamma(\alpha)} \left(\frac{\beta}{2 + \beta} \right)^\alpha$$

References

Xie, Y., J.J. Allaire, and G. Golemund. 2018. *R Markdown: The Definitive Guide*. CRC Press.