

EEB 603 – Chapter 1: Using the R Markdown language to promote Reproducible Science

Part B: Tables, Figures and References

Sven Buerki

Fall 2020

Contents

Learning outcomes	3
Associated files supporting this tutorial	3
Insert tables and figures in R Markdown document	3
Tables	3
Step-by-step protocol	4
Figures	5
Step-by-step protocol	5
Questions and tasks	6
Cross-referencing tables and figures	6
Cross-referencing tables	6
Cross-referencing figures	7
Practice cross-referencing tables and figures	7
Citing references in text and adding a bibliography/references section	7
The bibliography file	7
Procedure to do prior to citing references in an R Markdown document	7
Specifying a bibliography in the R Markdown file	7
The BibTeX format	8
Citation identifier: Connecting your .bib file with your .Rmd document	8
Citing references	8
Practice citing references	9
Adding a bibliography/references section	9

Formatting citations to journal styles	9
What is the citation style language (CSL)?	9
CSL repositories	10
How to use a CSL file in an R Markdown to format citations and bibliography?	10
References	10

Learning outcomes

This tutorial is devoted to part B of chapter 1 and provides students with opportunities to learn procedures to:

- Insert tables and figures in R Markdown document.
- Cross-reference tables and figures in the text.
- Cite references in the text and add a bibliography/references section.
- Format citations to journal styles.

Associated files supporting this tutorial

Although less used in part A, a set of files are provided to support teaching of material presented in this chapter. These files are deposited in the shared Google Drive at this path:

- `Reproducible_Science/Chapters/Chapter_1/Tutorial_files`

Files are as follows:

- `EEB603_Syllabus_BUERKI.Rmd`: This is the .Rmd file used to compile the syllabus of this class. This file provides a good source of information for the syntax and protocols described in this tutorial.
- `Bibliography_Reproducible_Science_2.bib`: This file contains references cited in **BibTeX** format.
- `AmJBot.csl`: This citation style language (CSL) file allows formatting citations and bibliography following citation style of *American Journal of Botany*.
- `Bioinformatic_workflow_PART2.pdf`: A pdf file containing the bioinformatic workflow taught in this class. This file will be used to learn how to incorporate a figure into R Markdown file.

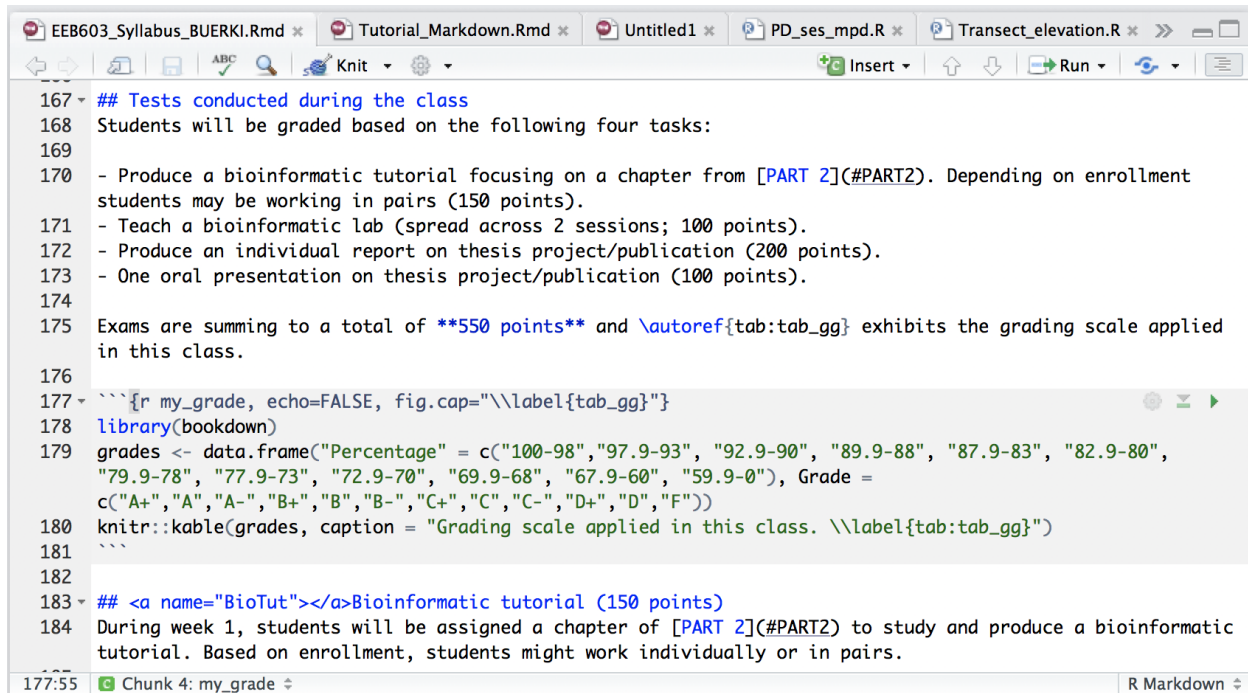
Insert tables and figures in R Markdown document

Tables

There will be more details about tables in chapter 9; however this tutorial introduces key concepts related to table making in R Markdown, more specifically on the following topics:

- Creating a table in R (the grading scale provided in the Syllabus; **Figure 1**).
- Assigning a table caption.
- Assigning a unique label to the table allowing further cross-referencing in the text.
- Displaying the table in the document.
- Cross-referencing the table in the text.

The first four topics are detailed below, whereas the last objective will be discussed in the next section



```



167 ## Tests conducted during the class
168 Students will be graded based on the following four tasks:
169
170 - Produce a bioinformatic tutorial focusing on a chapter from [PART 2](#PART2). Depending on enrollment
    students may be working in pairs (150 points).
171 - Teach a bioinformatic lab (spread across 2 sessions; 100 points).
172 - Produce an individual report on thesis project/publication (200 points).
173 - One oral presentation on thesis project/publication (100 points).
174
175 Exams are summing to a total of 550 points and \autoref{tab:tab_gg} exhibits the grading scale applied
    in this class.
176
177 ```{r my_grade, echo=FALSE, fig.cap="\label{tab_gg}`}
178 library(bookdown)
179 grades <- data.frame("Percentage" = c("100-98", "97.9-93", "92.9-90", "89.9-88", "87.9-83", "82.9-80",
    "79.9-78", "77.9-73", "72.9-70", "69.9-68", "67.9-60", "59.9-0"), Grade =
    c("A+", "A", "A-", "B+", "B", "B-", "C+", "C", "C-", "D+", "D", "F"))
180 knitr::kable(grades, caption = "Grading scale applied in this class. \label{tab:tab_gg}")
181 ```
182
183 ## <a name="BioTut"></a>Bioinformatic tutorial (150 points)
184 During week 1, students will be assigned a chapter of [PART 2](#PART2) to study and produce a bioinformatic
    tutorial. Based on enrollment, students might work individually or in pairs.
177:55 Chunk 4: my_grade

```

Figure 1: Snapshot of the section of the Rmd Syllabus highlighting the R code used to create the table of the grading scale applied in this class. The code includes the syntax used to create a table caption as well as the procedure to label it for further cross-referencing.

Step-by-step protocol

Please find below the steps required to:

1. Create and save a new blank .Rmd document using procedure learned in part A (using the pdf output format). This document will be used throughout this tutorial.
2. Add a first-level header entitled **Tables**.
3. Insert an R code chunk under your header by clicking on the **Insert** button  in the editor toolbar (Figure 1).
4. Write the R code displayed in Figure 1 into the code chunk section to generate the grading scale. Note that the `knitr::kable` function is applied to display the table in the document.
5. Test your code to check that it produces the expected table (using the **Run** button).
6. Edit the code chunk options line by adding a unique code chunk ID. This argument is displayed just after `{r`. Please name this code chunk `my_grade`.
7. Add a table caption by editing the `knitr::kable` function as shown in Figure 1.
8. Add a unique label (here `\label{tab:tab_gg}`) to allow further cross-referencing. This is done by amending the `knitr::kable` function. Notice: The `\label{}` function invokes the *bookdown* R package. This package should be installed on your computer prior to executing this code.
9. Knit your document using the **Knit** button  on the editor toolbar (Figure 1). The table should be automatically inserted into the document together with its caption.

Figures

There will be more details about figures in chapter 10; however this tutorial introduces key concepts related to figure making in R Markdown, more specifically on the following topics:

- Creating a figure in R (based on the cars dataset; [Figure 2](#)).
- Assigning a figure caption.
- Providing a unique label to the R code chunk allowing further cross-referencing in the text.
- Displaying the figure in the document.
- Cross-referencing the figure using its unique label.

The first four topics are detailed below, whereas the last objective will be discussed in the next section.

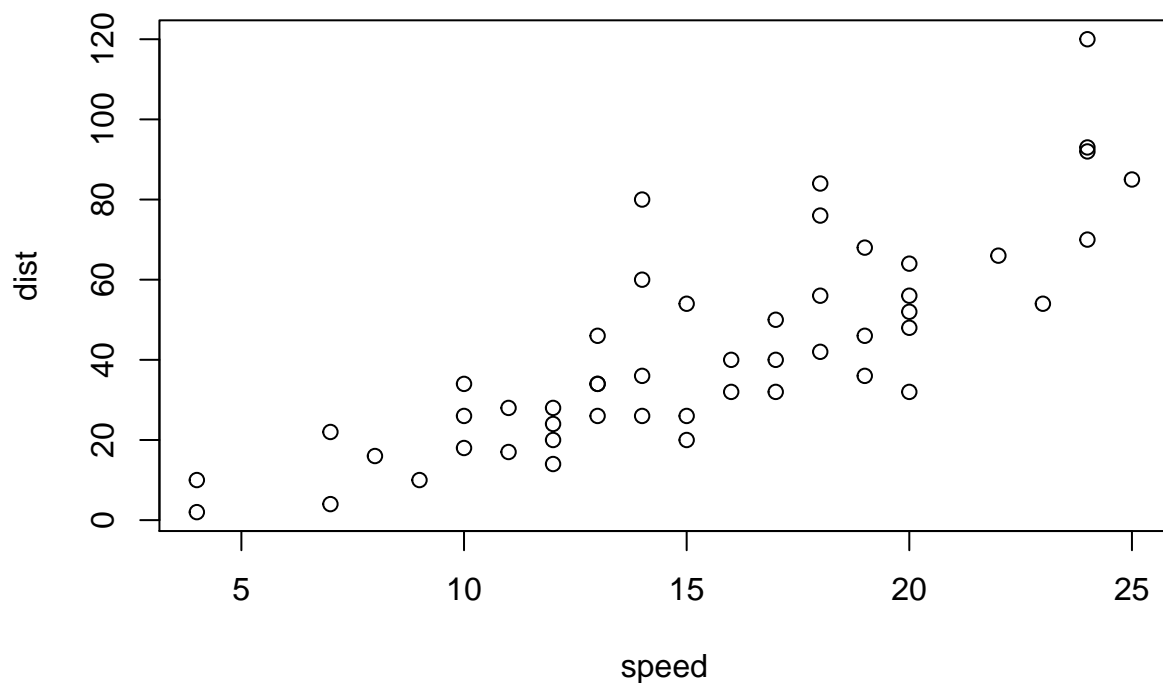



Figure 2: Plot of cars' speed in relation to distance.

Step-by-step protocol


Here, you will learn the R Markdown syntax and R code required to replicate [Figure 2](#):

1. Use the same `.Rmd` document as above to practice working with figures.
2. Add a first-level header entitled **Figures**.

3. Insert an R code chunk under your header by clicking on the **Insert** button  in the editor toolbar ([Figure 1](#)).

4. Write the following R code in the code chunk:

```
summary(cars)
plot(cars)
```

5. Edit the code chunk options line by adding the following arguments (each argument should be separated by a comma):
 - `echo=FALSE`
 - `results="hide"`
 - `fig.cap="Plot of cars' speed in relation to distance."`
 - `out.width = '100%'`
6. Add the unique label `cars` in the chunk options line (just after `{r}`) to enable further cross-referencing.
7. Test your code to check that it produces the expected plot (using the **Run** button).
8. Knit your document using the Knit button  **Knit** on the editor toolbar (Figure 1).

Questions and tasks

- What is the purpose of the `results="hide"` argument declared in the code chunk options line?
- Develop an R code using `knitr` to incorporate a figure into the R Markdown file based on an external file. Use your favorite figure (in `.jpeg` or `pdf`) to develop your code. **Tip:** Open `EEB603_Syllabus_BUERKI.Rmd` and look at the `part2` code chunk to learn more about the function to execute this task.

Cross-referencing tables and figures

Cross-referencing tables and figures in the text is now as easy as pie since labels have been added to these latter items. Cross-referencing works only when the document is compiled into `pdf` format since this procedure uses the `\autoref{}` TeX function. Tables and figures will be automatically numbered during the compilation process (one less thing to worry about).

Cross-referencing tables

To cross-reference a table in the R Markdown document use the following TeX syntax (see Figure 1 for an example focused on cross-referencing the grading scale table):

```
#General syntax
\autoref{tab:tab_ID}
tab_ID = Unique label referring to the table

#Example of syntax to refer to the grading scale
\autoref{tab:tab_gg}
```

This example is using the syntax from the *bookmark* R package (see Figure 1).

Cross-referencing figures

To cross-reference a figure in the R Markdown document use the following TeX syntax:

```
#General syntax
\autoref{fig:fig_ID}
fig_ID = Unique label referring to the figure in the R code chunk options line

#Example of syntax to refer to the plot above
\autoref{fig:cars}
```

Practice cross-referencing tables and figures

Students have to use their .Rmd document to practice cross-referencing tables and figures using procedures described above. To clearly define where you practice these new skills, please do so under a **Cross-referencing tables and figures** header.

Citing references in text and adding a bibliography/references section

The bibliography file

To cite references in the R Markdown document those have to be saved in a bibliography file using the BibTeX format. Other formats can be used, but the BibTeX format is open-source and easy to edit. Please see this webpage for more details on other formats: https://rmarkdown.rstudio.com/authoring_bibliographies_and_citations.html

Most journals allow saving citation of publications directly in BibTeX format, but when this feature is not available formats can be converted using online services (e.g. EndNote to BibTeX: <https://www.bruot.org/ris2bib/>).

Procedure to do prior to citing references in an R Markdown document

1. Save all your BibTeX references in a text file and make sure to add the .bib extension.
2. This file has to be deposited in the same folder as your .Rmd file.
3. You will be inserting the name of your references file in the YAML metadata section (see [below](#)).
4. You can visit this [webpage](#) and click on the Cite icon to download a citation in .bibtex format. More details on the BibTeX format is provided below.
5. References formatted in BibTeX format are available in associated file:
 - Bibliography_Reproducible_Science_2.bib.

Specifying a bibliography in the R Markdown file

The Pandoc program can automatically generate citations in the text and a bibliography/references section following various journal styles. In order to use this feature, you need to specify a bibliography file in the YAML metadata section (i.e. the first section of an R Markdown file). This can be done as follows:

```

---
title: "Sample Document"
output: pdf_document
bibliography: bibliography.bib
---

```

In the example above, “bibliography.bib” refers to the name of the file containing the BibTeX references.
Warning: this file has to be stored in the same directory as the .Rmd file.

The BibTeX format

Please find below an example of a reference formatted in BibTeX format:

```

# Example of BibTeX format for Baker (2016) published in Nature
@Article{Baker_2016,
  doi = {10.1038/533452a},
  url = {https://doi.org/10.1038/533452a},
  year = {2016},
  month = {may},
  publisher = {Springer Nature},
  volume = {533},
  number = {7604},
  pages = {452--454},
  author = {Monya Baker},
  title = {1,500 scientists lift the lid on reproducibility},
  journal = {Nature},
}

```

Citation identifier: Connecting your .bib file with your .Rmd document

The unique citation identifier of a reference (Baker_2016 in the example above) is set by the user in the BibTeX citation (see first line in the example provided above). This unique identifier is used to refer to the reference/publication in the R Markdown document and also allows citing references and generating the bibliography/references section.

Citing references

Citations go inside square brackets ([]) and are separated by semicolons. Each citation must have a key, composed of @ + the citation identifier (see above) as stored into the BibTeX file.

Please find below some examples on citation protocols:

```

#Syntax
Blah blah [see @Baker_2016, pp. 33-35; also @Smith2016, ch. 1].
Blah blah [@Baker_2016; @Smith2016].

```

Once knitted (using the  Knit button), the above code/syntax turns into:

Blah blah (see Baker, 2016 pp. 33–35; also Smith et al., 2016, ch. 1).

Blah blah (Baker, 2016; Smith et al., 2016).

A minus sign (-) before the @ will suppress mention of the author in the citation. This can be useful when the author is already mentioned in the text:

```
#Syntax  
Baker says blah blah [-@Baker_2016].
```

Once knitted, the above code/syntax turns into:

Baker says blah blah (2016).

You can also write an in-text citation, as follows:

```
#Syntax  
@Baker_2016 says blah.  
@Baker_2016 [p. 1] says blah.
```

Once knitted, the above code/syntax turns into:

Baker (2016) says blah.

Baker (2016 p. 1) says blah.

Practice citing references

Students have to use their .Rmd document to practice citing references in the text using procedures described above. To clearly define where you practice citing references, please do so under a **Citing references** header.

Adding a bibliography/references section

Upon knitting, a bibliography/references section will be automatically generated at the end of the document. Usually, we recommend adding a References header just after the last paragraph of the document as displayed below:

```
last paragraph...
```

```
# References
```

The bibliography will be inserted after this header (please see References section of this tutorial for more details).

Formatting citations to journal styles

In this section, we are studying how your bibliography can be automatically formatted following a journal style. This is achieved by providing the name of a citation style file (containing the protocol to format citations and bibliography following a journal style) in the YAML metadata section.

What is the citation style language (CSL)?

The **Citation Style Language** (CSL) was developed by an open-source project and aims at facilitating scholarly publishing by automating the formatting of citations and bibliographies. This project has developed the CSL and maintains a crowd sourced repository with over 8000 free CSL citation styles. Please see the following website for more details: <https://citationstyles.org>

CSL repositories

There are two main CSL repositories:


- **GitHub Repository:** <https://github.com/citation-style-language/styles>
- **Zotero Style Repository:** <https://www.zotero.org/styles>

How to use a CSL file in an R Markdown to format citations and bibliography?

Please follow the steps below to format your citations and bibliography following the citation style provided in a CSL file:

1. Download the CSL file using repositories provided above. Some journals provide their CSL files on their websites (one has been made available for you to use in the associated files: `AmJBot.csl`).
2. Save the CSL file in the same working directory as the `.Rmd` file.
3. Edit the YAML metadata section as follows to specify the CSL file:

```
# Add a "csl" argument and provide name of the CSL file (here AmJBot.csl)
---
title: "Sample Document"
output: pdf_document
csl: AmJBot.csl
bibliography: bibliography.bib
---
```

4. Knit the R Markdown document using the Knit  Knit button. The Pandoc program will use the information stored in the YAML metadata section to format the bibliography (citations and bibliography section) following the citation style provided in the CSL file. Do not forget to add a **References** header at the end of your `.Rmd` document.

References

BAKER, M. 2016. 1,500 scientists lift the lid on reproducibility. *Nature* 533: 452–454. Available at: <https://doi.org/10.1038/533452a>.

SMITH, J.F., T.H. PARKER, S. NAKAGAWA, J. GUREVITCH, ECOLOGY, and T.(. FOR T. IN EVOLUTION) WORKING GROUP. 2016. Promoting transparency in evolutionary biology and ecology. *Systematic Botany* 41: 495–497. Available at: <http://www.bioone.org/doi/abs/10.1600/036364416X692262>.