Data management, markdown, latex, notebooks (X. de Pedro) May 31 - http://seeds4c.org/180531

1.1. Introduction

1.1.1. Data Management

Using Tidyverse

- Cheatsheets:
  - https://www.rstudio.com/resources/cheatsheets/[1]
- Installation:

```
install.packages("tidyverse")
```

- Book:

> *R for Data Science. Garrett Grolemund, Hadley Wickham. O'Reilly January 2017 First Edition http://r4ds.had.co.nz*[2]
>
> *Welcome*
>
> *This is the website for "R for Data Science". This book will teach you how to do data science with R: You'll learn how to get your data into R, get it into the most useful structure, transform it, visualise it and model it. In this book, you will find a practicum of skills for data science. Just as a chemist learns how to clean test tubes and stock a lab, you'll learn how to clean data and draw plots—and many other things besides. These are the skills that allow data science to happen,*

> *and here you will find the best practices for doing each of these things with R. You'll learn how to use the grammar of graphics, literate programming, and reproducible research to save time. You'll also learn how to manage cognitive resources to facilitate discoveries when wrangling, visualising, and exploring data.*

- Exercises

See:

- https://www.tidyverse.org[3]

Using Radiant (Shiny app)

See also Radiant as GUI to allow business analytics using R and Shiny http://vnijs.github.io/radiant/[4]

Radiant is an open-source platform-independent browser-based interface for business analytics in R.

Key features

- Explore: Quickly and easily summarize, visualize, and analyze your data
- Cross-platform: It runs in a browser on Windows, Mac, and Linux
- Reproducible: Recreate results and share work with others as a state-file or an Rmarkdown report
- Programming: Integrate Radiant's analysis functions with your own R-code
- Context: Data and examples focus on business applications

1.1.2. Markdown

1.1.2.1. Short demo

RStudio

```
---
title: "Untitled"
author: "Ajuntament de Barcelona"
```

````
date: "28 febrer de 2018"
output: html_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>.

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 cars}
summary(cars)
```

## Including Plots

You can also embed plots, for example:

```{r pressure, 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.
````

1.1.2.2. Custom Example

See a custom example fetching and processing Public Health Data from Barcelona out of a PDF report with R:
http://rpubs.com/Xavi/ubiss16d3[5]

1.1.2.3. Lessons from RStudio

See all lessons from RStudio:

1. Introduction - https://rmarkdown.rstudio.com/lesson-1.html[6]
2. How It Works - https://rmarkdown.rstudio.com/lesson-2.html[7]
3. Code Chunks - https://rmarkdown.rstudio.com/lesson-3.html[8]

4. Inline Code - https://rmarkdown.rstudio.com/lesson-4.html[9]
5. Code Languages - https://rmarkdown.rstudio.com/lesson-5.html[10]
6. Parameters - https://rmarkdown.rstudio.com/lesson-6.html[11]
7. Tables - https://rmarkdown.rstudio.com/lesson-7.html[12]
8. Markdown Basics - https://rmarkdown.rstudio.com/lesson-8.html[13]
9. Output Formats - https://rmarkdown.rstudio.com/lesson-9.html[14]
10. Notebooks - https://rmarkdown.rstudio.com/lesson-10.html[15]
11. Slide Presentations - https://rmarkdown.rstudio.com/lesson-11.html[16]
12. Dashboards - https://rmarkdown.rstudio.com/lesson-12.html[17]
13. Websites - https://rmarkdown.rstudio.com/lesson-13.html[18]
14. Interactive Documents - https://rmarkdown.rstudio.com/lesson-14.html[19]
15. Cheatsheets - https://rmarkdown.rstudio.com/lesson-15.html[20]

1.1.3. Latex

The minimum content you need to include for LaTeX to build your document are a specification of the document class, indicators of document beginning and ending, and a little content. Comments in LaTeX are indicated with a leading “%” symbol.

You could also use other programs to write latex, such Texstudio (multiplatform, open source), among others.

See:

- https://support.rstudio.com/hc/en-us/articles/200552056-Using-Sweave-and-knitr[21]
- https://rpubs.com/YaRrr/SweaveIntro[22]
- http://keziamanlove.com/latex-in-rstudio/[23]

Example:



```
% trial .tex file %
\documentclass[10pt]{article}  % specifies document class (article) and point size
(10pt)

\begin{document}               % starts document

\title{Example Document}       % specifies big, fancy title
\maketitle                     % constructs big, fancy title
\section{Section 1}            % makes a section header
Here is some text in section 1.  Section 1 also contains an itemized list:
  \begin{itemize}              % initiates an itemized list
    \item Here is an item in the list
    \item Here is a second item
```

```
  \end{itemize}                % ends itemized list

\section{Section 2}            % makes header for section 2
Some text on section 2 here.  Section 2 contains an enumerated list.
  \begin{enumerate}            % initiates enumerated list
    \item A first enumerated item
    \item A second enumerated item
  \end{enumerate}              % ends enumerated list

\end{document}                 % ends document
```

Once you have this script written into your .tex file in RStudio, hit the Compile .pdf button at the upper right of the script window. Doing so will prompt LaTeX (via RStudio) to compiled the .pdf document you've specified. It will store the .pdf file (along with several auxiliary files LaTeX files) in your working directory

Latex Collaborative Editors online

Overleaf

Overleaf (https://www.overleaf.com[24]) is yet another online LaTeX editor. Similar to ShareLaTeX, it offers separate pricing plans for professionals and students. It also includes a free plan where you can sync with GitHub, check your revision history, and add multiple collaborators.

There's a limit on the number of files you can create per project – so it could bother if you are a professional working with LaTeX documents most of the time.

ShareLatex

ShareLaTeX (https://www.sharelatex.com[25]) is an online LaTeX editor. If you want someone (or a group of people) to collaborate on documents you are working on, this is what you need.

It offers a free plan along with several paid packages. Even the students of Harvard University & Oxford University utilize this for their projects. With the free plan, you get the ability to add one collaborator.

The paid packages let you sync the documents on GitHub and Dropbox along with the ability to record the full document history. You can choose to have multiple collaborators as per your plan. For students, there's a separate pricing plan available.

1.1.3.1. Learn LaTeX in 30 minutes

From sharelatex.com:
https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes[26]

You can practice online here: https://latexbase.com[27]

- 1 What is LaTeX?[28]
- 2 Why learn LaTeX?[29]
- 3 Writing your first piece of LaTeX[30]
- 4 The preamble of a document[31]
- 5 Adding a title, author and date[32]
- 6 Adding comments[33]
- 7 Bold, italics and underlining[34]
- 8 Adding images[35]
    - 8.1 Captions, labels and references[36]
- 9 Creating lists in LaTeX[37]
    - 9.1 Unordered lists[38]
    - 9.2 Ordered lists[39]
- 10 Adding math to LaTeX[40]
- 11 Basic Formatting[41]
    - 11.1 Abstracts[42]
    - 11.2 Paragraphs and newlines[43]
    - 11.3 Chapters and Sections[44]
- 12 Creating tables[45]
    - 12.1 Creating a simple table in LaTeX[46]
    - 12.2 Adding borders[47]
    - 12.3 Captions, labels and references[48]
- 13 Adding a Table of Contents[49]
- 14 Downloading your finished document[50]

1.1.4. Notebooks

You could use either R Notebook (using R Markdown from within RStudio) or Jupyter notebook, both of them allow you to embed R o rpython code inside, even if the first notebook is mainly coming from the R's community, and the second, from Python's community. See a few comparisons here:

- https://www.datacamp.com/community/blog/jupyter-notebook-r[51]
- http://minimaxir.com/2017/06/r-notebooks/[52]

We will stick in this session to the R Notebook from RStudio, while in further sessions in this course you might get introduced to Python Jupyter's notebook.



```
---
title: "R Notebook"
output: html_notebook
---
```

```
This is an [R Markdown](http://rmarkdown.rstudio.com) Notebook. When you execute
code within the notebook, the results appear beneath the code.

Try executing this chunk by clicking the *Run* button within the chunk or by
placing your cursor inside it and pressing *Ctrl+Shift+Enter*.

```{r}
plot(cars)
```

Add a new chunk by clicking the *Insert Chunk* button on the toolbar or by pressing
Ctrl+Alt+I.

When you save the notebook, an HTML file containing the code and output will be
saved alongside it (click the *Preview* button or press *Ctrl+Shift+K* to preview
the HTML file).

The preview shows you a rendered HTML copy of the contents of the editor.
Consequently, unlike *Knit*, *Preview* does not run any R code chunks. Instead, the
output of the chunk when it was last run in the editor is displayed.
```

1.2. Hands-on work

1.2.1. Hello-World Demo

Follow this tutorial as a starting point:
https://medium.com/@sorenlind/create-pdf-reports-using-r-r-markdown-latex-and-knitr-on-windows-10-952b0c48bfa9[53]

1.2.2. A step-by-step markdown example

A step by step guide to learn the basics of RMarkdown
https://gist.github.com/jeromyanglim/2716336[54]

1.2.3. An R Markdown Template for Academic Manuscripts

"You should consider no longer using LaTeX as a front-end for your manuscripts. Use a wrapper for LaTeX instead, like R Markdown...":
http://svmiller.com/blog/2016/02/svm-r-markdown-manuscript/[55]

1.2.4. Extra tips: selected Rstudio Addins

Addins List: install with:



```
install.packages("addinslist")
```

Recommended Rstudio addins for this course:

- CRANsearcher
  to search in package names and descriptions at once
- GitGadget
  to easily manage gitlab repo collaborative work from within RStudio
- Radiant
  to easily do plenty of Data Management from within a simple yet powerful GUI
- Remedy
  to easily embed rmarkdown tags in context where/when needed
- Rhandontable
  to visually edit a dataframe

Alias names for this page:
markdown | rmarkdown | latex | notebooks | 180531

---

[1] https://www.rstudio.com/resources/cheatsheets/
[2] http://r4ds.had.co.nz
[3] https://www.tidyverse.org
[4] http://vnijs.github.io/radiant/
[5] http://rpubs.com/Xavi/ubiss16d3
[6] https://rmarkdown.rstudio.com/lesson-1.html
[7] https://rmarkdown.rstudio.com/lesson-2.html
[8] https://rmarkdown.rstudio.com/lesson-3.html
[9] https://rmarkdown.rstudio.com/lesson-4.html
[10] https://rmarkdown.rstudio.com/lesson-5.html
[11] https://rmarkdown.rstudio.com/lesson-6.html
[12] https://rmarkdown.rstudio.com/lesson-7.html
[13] https://rmarkdown.rstudio.com/lesson-8.html
[14] https://rmarkdown.rstudio.com/lesson-9.html
[15] https://rmarkdown.rstudio.com/lesson-10.html
[16] https://rmarkdown.rstudio.com/lesson-11.html
[17] https://rmarkdown.rstudio.com/lesson-12.html
[18] https://rmarkdown.rstudio.com/lesson-13.html
[19] https://rmarkdown.rstudio.com/lesson-14.html
[20] https://rmarkdown.rstudio.com/lesson-15.html
[21] https://support.rstudio.com/hc/en-us/articles/200552056-Using-Sweave-and-knitr
[22] https://rpubs.com/YaRrr/SweaveIntro

[23] http://keziamanlove.com/latex-in-rstudio/

[24] https://www.overleaf.com

[25] https://www.sharelatex.com

[26] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes

[27] https://latexbase.com

[28] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#What_is_LaTeX.3F

[29] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Why_learn_LaTeX.3F

[30] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Writing_your_first_piece_of_LaTeX

[31] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#The_preamble_of_a_document

[32] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Adding_a_title.2C_author_and_date

[33] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Adding_comments

[34] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Bold.2C_italics_and_underlining

[35] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Adding_images

[36] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Captions.2C_labels_and_references

[37] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Creating_lists_in_LaTeX

[38] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Unordered_lists

[39] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Ordered_lists

[40] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Adding_math_to_LaTeX

[41] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Basic_Formatting

[42] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Abstracts

[43] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Paragraphs_and_newlines

[44] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Chapters_and_Sections

[45] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Creating_tables

[46] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Creating_a_simple_table_in_LaTeX

[47] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Adding_borders

[48] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Captions.2C_labels_and_references_2

[49] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Adding_a_Table_of_Contents

[50] https://es.sharelatex.com/learn/Learn_LaTeX_in_30_minutes#Downloading_your_finished_document

[51] https://www.datacamp.com/community/blog/jupyter-notebook-r

[52] http://minimaxir.com/2017/06/r-notebooks/

[53] https://medium.com/@sorenlind/create-pdf-reports-using-r-r-markdown-latex-and-knitr-on-windows-10-952b0c48bfa9

[54] https://gist.github.com/jeromyanglim/2716336

[55] http://svmiller.com/blog/2016/02/svm-r-markdown-manuscript/