Hiroshi Suzuki

2023-02-27

Contents

AI

2019

2019 9

2020 Data Analysis for Researchers 20

PDF ePub

PDF ePub

- PDF
- ePub

Chapter 1

R python R R Studio IDE integrated development environment, R Posit Cloud Reproducibility Literate Programming R Markdown R Markdown bookdown Bookdown R

APPENDIX **??**

variable object

Appendix A

MOOCs

2019 9 17

A.1 YouTube

PDF

A.2

A.2.1

A.2.1.1

A.2.1.1.1

- 5 :2016 1 22 [1]
  - https://www8.cao.go.jp/cstp/kihonkeikaku/index5.html
- :2016 12 21 -

```
1.          ( )
2.
3.
https://www8.cao.go.jp/cstp/kihonkeikaku/index5.html

-
```

[1] 1 2

- (6)

- (D-DRIVE[2])2017 ~

- – 20 :2019 1 8
  – http://www.mext.go.jp/b_menu/shingi/chousa/koutou/095/gaiyou/1412367.htm

- : 2019 3 26
  –
  – https://www.meti.go.jp/shingikai/economy/risukei_jinzai/20190326_report.html

- ( )
  – AI :2019 3 29 https://www8.cao.go.jp/cstp/aigensoku.pdf
  – AI 2019 AI 2019 6 11 https://www.kantei.go.jp/jp/singi/tougou-innovation/pdf/aisenryaku2019.pdf

1. 1: (50 /) AI [MOOC]
2. 2: (100 /) AI
3. 3: (AI)

A.2.1.2

http://www.mi.u-tokyo.ac.jp/consortium/index.html

-
-

A.2.1.3

- 2019

A.2.2

A.2.2.1 IR

IR IR

A.2.2.2

2019 3 MOOCs

[2]Doctoral program for Data-Related InnoVation Expert

A.2.2.3 : (Data Analysis for Researchers)

- ( ) 3
- :Rotary Peace Fellow, The Project for Human Resource Development Scholarship (JDS) 4 10-25
- :
- :2 (70 × 2 × 10) 1 1
- : (R Markdown etc.)
- :2014-2015[3], (2016), 2017

A.2.2.4

1. Introduction to R, Open Data and Free Software 2. Basic R Objects and Commands
2. Data Frame Manipulation
3. Linear Regression and Graphics
4. Dynamic Documents Using Rmarkdown 6. Statistical analysis with R II
5. Statistical analysis with R III
6. Statistical analysis with R IV
7. Guest Lecture and preparation for presentations
8. Final presentations

A.2.2.5

R Studio (PC) R R Markdown

- Base R
  - cars:
  - iris:
- package MASS
- WDI: World Bank Development indicators for R

```
library(WDI)
#GDP (current US$)
gdp <- WDI(country = c("US", "JP", "CN", "KR"),
indicator = "NY.GDP.MKTP.CD",
start = 1960, end = 2017)
```

  - wbstats
- Quandl package: https://www.quandl.com/tools/r
- Google Trends: https://trends.google.co.jp/
- Yahoo Finance: https://finance.yahoo.com/quote/DATA/

[3]32015 (Computing for Researchers)

A.2.3

A.2.3.1

? ?

- :
- AI
- ( )

A.2.3.1.1

- ( )

A.2.3.2 ( )

- ( )
  - ( ) ? ?
  - ( ) AI
  - ( )

A.2.3.3

? Liberal Arts?

A.2.3.4

Teaching to Learning

- *
  - - ( ) Community of Inquiry (CoI), - Communication of Facts

A.2.3.5

Students with Various Backgrounds

- :
- ( )

A.2.3.6 Literacy

Expand your horizon!

- :
  -
  - Q & A ( )
  -
- :
  - I, II, A, B - :
  -
- :
  -
  -

A.2.3.7 Resources

IT / Cloud

- (Open/Public Data)
- (Online/Cloud)
- (Free and online/cloud system)

A.2.3.8 :

1.
2.
3. http://www.soumu.go.jp/menu_seisaku/ictseisaku/ictriyou/opendata/

A.2.3.9 World Bank: Open Data Defined

The term “Open Data” has a very precise meaning. Data or content is open if anyone is free to use, re-use or redistribute it, subject at most to measures that

preserve provenance and openness.

1. The data must be **legally open**, which means they must be placed in the public domain or under liberal terms of use with minimal restrictions.

2. The data must be **technically open**, which means they must be published in electronic formats that are machine readable and non-proprietary, so that anyone can access and use the data using common, freely available software tools. Data must also be publicly available and accessible on a public server, without password or firewall restrictions. To make Open Data easier to find, most organizations create and manage Open Data catalogs.

A.2.3.10

List of Open Data Catalogue

- :https://www.data.go.jp
- :

  1.
  2.
  3.

  - :https://www.data.go.jp/list-of-database/
  - :https://www.jma.go.jp/jma/menu/menureport.html

- U.S. Government's Open Data: https://www.data.gov
- EU Open Data Portal: http://data.europa.eu/euodp/en/home
- UK Open Data: https://data.gov.uk
- World Bank: New Ways of Looking at Poverty
  - Open Data: https://data.worldbank.org
  - World Development Indicators: http://datatopics.worldbank.org/world-development-indicators/
- UN Data: http://data.un.org
- WHO Data: https://www.who.int/gho/en/
- Google Public Data: :7 :136 https://www.google.com/publicdata/directory
- Open Knowledge Foundation: https://okfn.org
  - Global Open Data Index: https://index.okfn.org

A.2.3.10.1 Free Software, Online Access R

- R Project for Statistical Computing: https://www.r-project.org
- R Studio: https://www.rstudio.com
- R Studio Cloud: https://rstudio.cloud

A.2.3.10.2 Python

- Phython: https://www.python.org
- Anaconda: https://www.anaconda.com
- Jupyter Notebook Cloud: Binder, Kaggle Kernels, Google Collaborate, CoCalc, PaizaCloud, etc.

A.2.3.10.3 Free Software

- Free Software, Free Society: Selected Essays of Richard M. Stallman: ht tps://www.gnu.org/philosophy/ fsfs/rms-essays.pdf
- Richard Stallman TEDxGeneva 2014: https://youtu.be/Ag1AKIl_2GM

A.2.3.10.4 Online Learning Source

List of Online Help and Mini Courses

- Online
  - TutorialPoint: https://www.tutorialspoint.com/
  - DataCamp: https://www.datacamp.com/home
  - Code Academy: https://www.codecademy.com
  - RStudio Premier: https://rstudio.cloud/learn/primers
- User Community

A.2.3.10.5 MOOCs

- OED: MOOC n. massive open online course, an educational course made available to a large number of people via the internet.
- First MOOC: 2008 by Dave Cormier, Connectivism and Connective Knowledge (CCK08)
- MIT OpenCourseWare 2002: ( )OER Stanford U Model: UC Berkeley:
- MOOC : 2012 Stanford U. MIT

A.2.4 Massive Open Online Courses (MOOCs)

Moocs

A.2.4.1 MOOCs **

-
-
-
- (datasets) ( ) *

A.2.4.2 MOOCs

-
-
-
-

A.2.4.3 Moocs

A.2.4.3.1 Coursera Stanford U.

- Stanford U.: Machine Learning https://www.coursera.org/learn/machine-learning
- (:2,513,476)
- Johns Hopkins U.: Data Science, 10 courses [, R] https://www.coursera.org/specializations/jhu-data-science
- U. of Michigan: Applied Data Science, 5 courses [, python] https://www.coursera.org/specializations/data-science-python
- 100% U. of Illinois, U. of Michigan, U. of Colorado,

A.2.4.3.2 edX MIT Harvard U.

- Harvard U.: Data Science, 9 courses [, R]
- MIT: Statistics and Data Science, 5 courses [Graduate Level]
- https://www.edx.org/micromasters/mitx-statistics-and-data-science
- Microsoft, IBM, UC SanDiego

A.2.4.4 Professional Certificate in Data Science

A.2.4.4.1 : HarvardX, through edX

- URL: https://online-learning.harvard.edu/series/professional-certificate-data-science
- Book: https://rafalab.github.io/dsbook/ (R Markdown Document)

A.2.4.4.2 R Data Camp Assessment

```
1. Data Science: R Basics;      R
2. Data Science: Visualization;
3. Data Science: Probability;
4. Data Science: Inference and Modeling;
5. Data Science: Productivity Tools; Unix, Git, GitHub, R Markdown
```

```
6. Data Science: Wrangling;
7. Data Science: Linear Regression;
8. Data Science: Machine Learning;
9. Data Science: Capstone
```

A.2.4.4.3 Professional Certificate in Data Science

8 weeks 8 weeks 8 weeks 8 weeks 8 weeks 8 weeks 8 weeks 8 weeks 2 weeks Required R Packages for Examples: tidyverse, dslabs: https://cran.r-project.org/web/packages/ dslabs/dslabs.pdf

-
- 2010 FBI
- Gapminder: Almost nobody knows the basic global facts! (Gapminder Test)
  - TED (Hans Rosling) https://www.gapminder.org
  - Health and income outcomes for 184 countries from 1960 to 2016
  - Country, Year, Infant deaths per 1000, Life expectancy in years, Average of children per woman, Country population, GDP, Continent, Geographical region
- Brexit
- 2016
- UC Berkeley
-
-

A.2.4.4.4 JMOOC MOOC

(1)

- Week 1
  -
  - - gacco:
- Week 2
  -
  - :
- Week 3
  - Excel, R, Python
- Week 4
  -
  - -
  -

A.2.4.4.5 JMOOC: https: // www. jmooc. jp 10 8

- II: ( )

A.2.5

A.2.5.1

-
-
-
-
-
-

A.2.5.2 R Markdown Or jupyter notebook for python

- Reproducible Research:
- Literate Programming: Code, Script * html, doc, presentation, pdf (using TeX)
-

A.2.5.3

-
-
-
- Empirical ([,])
-

  Quote: On Listening to Lectures, by Plutarch The correct analogy for the mind is not a vessel that needs filling, but wood that needs igniting - no more - and then it motivates one towards originality and instills the desire for truth. (https://quoteinvestigator.com/2013/03/28/mind-fire/)

  ( )

A.2.6

A.2.6.1

Quote: Apple co-founders Steve Jobs and Steve Wozniak didn't have degrees when they launched what has become one of the most valuable companies in the world. And now Apple CEO Tim Cook is

spreading the word that would-be programmers really don't need the endorsement of a university to be able to create something of commercial value, such as an app for the Apple App Store. https://www.zdnet.com/article/apple-ceo-tim-cook-you-dont-need-a-degree-to-code-mobile-apps/

A.2.6.2

AI

Appendix B

MOOCs

3 12 (II) covid-19

B.1

AI [1] 2019 6 11 AI 2019 AI
( ) AI
2025

1. (50 /) AI [MOOC]
2. (100 /) AI
3. (AI)

6 25% 1500 AI (100 /) [1] 6
AI [3] 2000 [2] GP [3]
2012 [4] [4] STEAM
1, 3 ( )
AI AI

[1] 5 20 22 11 IT EXPO
[2] 22000.6 ( )
[3] 3
[4] 42012.8

B.2

AI 2019 AI
AI AI

AI (Artificial Intelligence( ) AI
2019 AI AI AI
AI

([11])

(Decision Science) (Empirical
Study) (,) ( )
AI
Google, Amazon, Netflix
(Recommendation System)

[5]
AI AI

([12])

AI
AI AI Artificial Gen-
eral Intelligence (AGI) AI
AI

B.3

B.3.1

[5] ([7, 9])

(IoT,
Big Data)

(Public Data)
[6]
(R, Python) edX, Coursera, JMOOC MOOCs

‘Skills and Self - ID Top Factors’
[13]
([8]) GAFA

B.3.2

TA TA

2025 5 AI

( )

R Free

( )

AI Respect

B.3.3

( )

( )

[6] 20

B.3.4

1. Broad Band Internet Access
   Broad Band Internet Access
2. RStudio (for R), Jupyter (for python) Data
   Science Interactive Security
3. Team and Collaborative Teaching
4. MOOCs OCW
   Interactive
5. Fact Base,
   Evidence Base AI Black Box
   Fake News Hate Speach
   Data Sciece r

Learning Goals

- A. Data Science: What? and Why?
  - Data Science ? AI ? Data Scientists?
  - Data Science ? -
- B. Business Data Science
- C. Data Science :Data Data
- D. Data Analysis

B.4 References

[1] AI (AI) https://www.kantei.go.jp/jp/singi/ai_senryaku/index.html

[2] AI 2019 AI 2019 6 11 https://www. kantei.go.jp/jp/singi/tougou-innovation/pdf/aisenryaku2019.pdf

[3] AI https: //www.kantei.go.jp/jp/singi/ai_senryaku/suuri_datascience_ai/dai1/sankou3.pdf

[4] https://www.mhlw.go.jp/toukei/ learning/index.html

[5]

[6] ~ Moocs (:)

[7] AUTOMATE THIS: how algorithm came to rule our world by Christopher Steiner

[8] Doing Data Science by Rachel Schutt and Cathy O'Neil

[9] Weapon's of Math Destruction by Cathy O'Neil AI

[10] edX Professional Certificate in Data Science: https://online-learning.harvard.edu/series/professional-certificate-data-science edX Data Science Ethics: https://www.edx.org/course/data-science-ethics

[11] Factfulness by Hans Rosling, et. al.

[12] How I learned to understand the world by Hans Rosling with Fanny Haergestam

[13] Analyzing the Analyzers (O'Reilly) by Harlan Harris, Sean Murphy, and Marck Vais- man

Appendix C

Bookdown

C.1 About

This is a *sample* book written in **Markdown**. You can use anything that Pandoc's Markdown supports; for example, a math equation $a^2 + b^2 = c^2$.

C.1.1 Usage

Each **bookdown** chapter is an .Rmd file, and each .Rmd file can contain one (and only one) chapter. A chapter *must* start with a first-level heading: `# A good chapter`, and can contain one (and only one) first-level heading.

Use second-level and higher headings within chapters like: `## A short section` or `### An even shorter section`.

The `index.Rmd` file is required, and is also your first book chapter. It will be the homepage when you render the book.

C.1.2 Render book

You can render the HTML version of this example book without changing anything:

1. Find the **Build** pane in the RStudio IDE, and
2. Click on **Build Book**, then select your output format, or select "All formats" if you'd like to use multiple formats from the same book source files.

Or build the book from the R console:

```
bookdown::render_book()
```

To render this example to PDF as a `bookdown::pdf_book`, you'll need to install XeLaTeX. You are recommended to install TinyTeX (which includes XeLaTeX): https://yihui.org/tinytex/.

C.1.3 Preview book

As you work, you may start a local server to live preview this HTML book. This preview will update as you edit the book when you save individual .Rmd files. You can start the server in a work session by using the RStudio add-in "Preview book", or from the R console:

```
bookdown::serve_book()
```

C.2 Hello bookdown

All chapters start with a first-level heading followed by your chapter title, like the line above. There should be only one first-level heading (`#`) per .Rmd file.

C.2.1 A section

All chapter sections start with a second-level (`##`) or higher heading followed by your section title, like the sections above and below here. You can have as many as you want within a chapter.

An unnumbered section

Chapters and sections are numbered by default. To un-number a heading, add a `{.unnumbered}` or the shorter `{-}` at the end of the heading, like in this section.

C.3 Cross-references

Cross-references make it easier for your readers to find and link to elements in your book.

C.3.1 Chapters and sub-chapters

There are two steps to cross-reference any heading:

1. Label the heading: `# Hello world {#nice-label}`.
    - Leave the label off if you like the automated heading generated based on your heading title: for example, `# Hello world` = `# Hello world {#hello-world}`.
    - To label an un-numbered heading, use: `# Hello world {-#nice-label}` or `{# Hello world .unnumbered}`.
2. Next, reference the labeled heading anywhere in the text using `\@ref(nice-label)`; for example, please see Chapter C.3.

- If you prefer text as the link instead of a numbered reference use: any text you want can go here.

C.3.2 Captioned figures and tables

Figures and tables *with captions* can also be cross-referenced from elsewhere in your book using `\@ref(fig:chunk-label)` and `\@ref(tab:chunk-label)`, respectively.

See Figure C.1.

```
par(mar = c(4, 4, .1, .1))
plot(pressure, type = 'b', pch = 19)
```

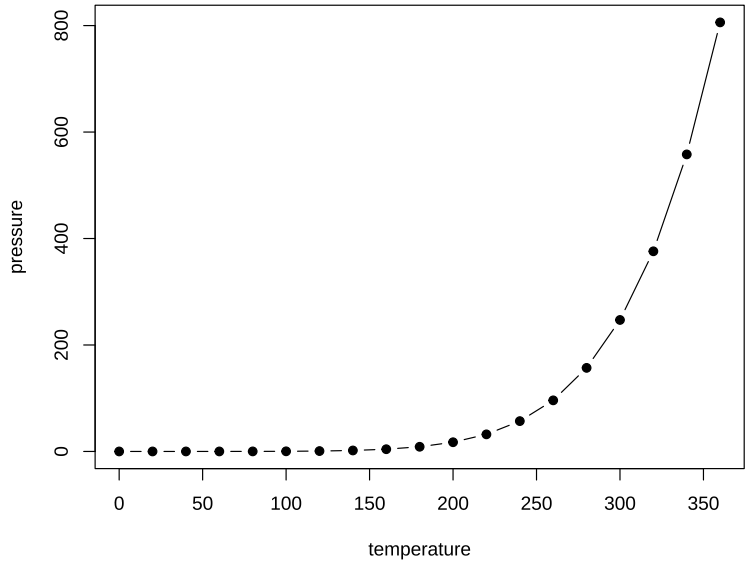

Figure C.1: Here is a nice figure!

Don't miss Table C.1.

```
knitr::kable(
  head(pressure, 10), caption = 'Here is a nice table!',
  booktabs = TRUE
)
```

C.4 Parts

You can add parts to organize one or more book chapters together. Parts can be inserted at the top of an .Rmd file, before the first-level chapter heading in that same file.

Table C.1: Here is a nice table!

temperature	pressure
0	0.0002
20	0.0012
40	0.0060
60	0.0300
80	0.0900
100	0.2700
120	0.7500
140	1.8500
160	4.2000
180	8.8000

Add a numbered part: `# (PART) Act one {-}` (followed by `# A chapter`)

Add an unnumbered part: `# (PART\*) Act one {-}` (followed by `# A chapter`)

Add an appendix as a special kind of un-numbered part: `# (APPENDIX) Other stuff {-}` (followed by `# A chapter`). Chapters in an appendix are prepended with letters instead of numbers.

C.5 Footnotes and citations

C.5.1 Footnotes

Footnotes are put inside the square brackets after a caret `^[]`. Like this one [1].

C.5.2 Citations

Reference items in your bibliography file(s) using `@key`.

For example, we are using the **bookdown** package (Xie, 2023) (check out the last code chunk in index.Rmd to see how this citation key was added) in this sample book, which was built on top of R Markdown and **knitr** (Xie, 2015) (this citation was added manually in an external file book.bib). Note that the `.bib` files need to be listed in the index.Rmd with the YAML `bibliography` key.

The `bs4_book` theme makes footnotes appear inline when you click on them. In this example book, we added `csl: chicago-fullnote-bibliography.csl` to the `index.Rmd` YAML, and include the `.csl` file. To download a new style, we recommend: https://www.zotero.org/styles/

[1] This is a footnote.

The RStudio Visual Markdown Editor can also make it easier to insert citations: https://rstudio.github.io/visual-markdown-editing/#/citations

C.6 Blocks

C.6.1 Equations

Here is an equation.

$$f(k) = \binom{n}{k} p^k (1-p)^{n-k} \tag{C.1}$$

You may refer to using `\@ref(eq:binom)`, like see Equation (C.1).

C.6.2 Theorems and proofs

Labeled theorems can be referenced in text using `\@ref(thm:tri)`, for example, check out this smart theorem C.1.

Theorem C.1. *For a right triangle, if c denotes the* length *of the hypotenuse and a and b denote the lengths of the* ***other*** *two sides, we have*

$$a^2 + b^2 = c^2$$

Read more here https://bookdown.org/yihui/bookdown/markdown-extensions-by-bookdown.html.

C.6.3 Callout blocks

The `bs4_book` theme also includes special callout blocks, like this `.rmdnote`.

You can use **markdown** inside a block.

```
head(beaver1, n = 5)
#>   day time  temp activ
#> 1 346  840 36.33     0
#> 2 346  850 36.34     0
#> 3 346  900 36.35     0
#> 4 346  910 36.42     0
#> 5 346  920 36.55     0
```

It is up to the user to define the appearance of these blocks for LaTeX output.

You may also use: `.rmdcaution`, `.rmdimportant`, `.rmdtip`, or `.rmdwarning` as the block name.

The R Markdown Cookbook provides more help on how to use custom blocks to design your own callouts: https://bookdown.org/yihui/rmarkdown-cookbook/custom-blocks.html

C.7 Sharing your book

C.7.1 Publishing

HTML books can be published online, see: https://bookdown.org/yihui/bookdown/publishing.html

C.7.2 404 pages

By default, users will be directed to a 404 page if they try to access a webpage that cannot be found. If you'd like to customize your 404 page instead of using the default, you may add either a `_404.Rmd` or `_404.md` file to your project root and use code and/or Markdown syntax.

C.7.3 Metadata for sharing

Bookdown HTML books will provide HTML metadata for social sharing on platforms like Twitter, Facebook, and LinkedIn, using information you provide in the `index.Rmd` YAML. To setup, set the `url` for your book and the path to your `cover-image` file. Your book's `title` and `description` are also used.

This `bs4_book` provides enhanced metadata for social sharing, so that each chapter shared will have a unique description, auto-generated based on the content.

Specify your book's source repository on GitHub as the `repo` in the `_output.yml` file, which allows users to view each chapter's source file or suggest an edit. Read more about the features of this output format here:

https://pkgs.rstudio.com/bookdown/reference/bs4_book.html

Or use:

```
?bookdown::bs4_book
```

Bibliography

Xie, Y. (2015). *Dynamic Documents with R and knitr*. Chapman and Hall/CRC, Boca Raton, Florida, 2nd edition. ISBN 978-1498716963.

Xie, Y. (2023). *bookdown: Authoring Books and Technical Documents with R Markdown*. R package version 0.32.