

Part 1. R 프로그래밍 (데이터 분석 전문가 양성과정)

10

분석보고서 작성

경북대학교 배준현 교수
(joonion@knu.ac.kr)



10. 분석보고서 작성

- 재현가능한 연구: *Reproducible Research*
 - 논문에서 주장하고자 하는 연구 결과를 누구나 재현할 수 있도록
 - 실제 연구에 사용한 데이터와 소스 코드를 공개 및 배포하려는 운동
 - R 커뮤니티에서 특히 활발하게 거론되는 연구 문화
 - R markdown 등으로 연구논문의 결과를 공개 및 배포할 수 있음



10. 분석보고서 작성

■ R *markdown*:

- **마크다운**: 서식이 있는 문서를 작성할 수 있는 간단한 문법
- R 마크다운: R 소스 코드와 마크다운 문서를 통합
 - 연구과정에 대한 **설명**(논문), **데이터**, **소스 코드**를 통합 배포
 - R 마크다운 문서의 확장자: **.Rmd**
 - **knitr**: HTML, PDF, Word 등의 문서로 변환





10. 분석보고서 작성

- R 마크다운 문서 생성: File > New File > R Markdown

The screenshot shows the 'New R Markdown' dialog box. On the left, there is a list of options: 'Document' (selected), 'Presentation', 'Shiny', and 'From Template'. On the right, there are input fields for 'Title' (filled with 'Reproducible Research with R'), 'Author' (filled with 'joonion'), and 'Date' (filled with '2022-05-23'). Below these fields is a checkbox labeled 'Use current date when rendering document' which is unchecked. Under the heading 'Default Output Format:', there are three radio button options: 'HTML' (unselected), 'PDF' (selected), and 'Word' (unselected). Below the 'PDF' option, there is a note: 'PDF output requires TeX (MiKTeX on Windows, MacTeX 2013+ on OS X, TeX Live 2013+ on Linux)'. Below the 'Word' option, there is a note: 'Previewing Word documents requires an installation of MS Word (or Libre/Open Office on Linux)'. At the bottom left is a button labeled 'Create Empty Document'. At the bottom right are 'OK' and 'Cancel' buttons.



10. 분석보고서 작성

■ YAML Header

- YAML: YAML Ain't Markup Language
 - Rmd 파일의 시작 부분에서 문서 형식을 설정하는 용도로 사용하는 포맷

```
---  
title: "Doing Reproducible Research with R"  
author: "Joonion Bae"  
date: '2022-05-23'  
output:  
  html_document:  
    df_print: paged  
  pdf_document: default  
  word_document: default  
---
```



10. 분석보고서 작성

- 마크다운 문법을 이용한 문서 작성 방법

```
## Writing with Markdown
```

```
Publish or perish! \
```

```
\
```

```
This is an important note to a researcher. \
```

```
### Publish or perish!
```

```
#### Publish or perish!
```

```
##### Publish or perish!
```

```
How to publish not to be perished?
```

```
1. Do your research.
```

```
2. Write your paper.
```

```
3. Submit to a journal
```

```
- If accepted, goto 1.
```

```
- If not accpted, goto 1.
```



10. 분석보고서 작성

- HTML 문서의 생성: RStudio 상단의 `knitr` > `Knit to HTML` 클릭

Doing Reproducible Research with R

Joonion Bae

2022-05-23

Writing with Markdown

Publish or perish!

This is an **important** note to a *researcher*.

Publish or perish!

Publish or perish!

Publish or perish!

How to publish not to be perished?

1. Do your research.
2. Write your paper.
3. Submit to a journal
 - If accepted, goto 1.
 - If not accepted, goto 1.



10. 분석보고서 작성

- Code Chunk: 마크다운 문서 내부에서 R 소스 코드를 포함하고 실행할 수 있음

```
```{r summary}
library(palmerpenguins)
library(tidyverse)
penguins %>%
 group_by(species) %>%
 summarise(across(where(is.numeric), mean, na.rm = T))
```
```




10. 분석보고서 작성

```
library(palmerpenguins)
library(tidyverse)
```

```
## — Attaching packages — tidyverse 1.3.1 —
```

```
## ✓ ggplot2 3.3.6    ✓ purrr 0.3.4
## ✓ tibble 3.1.7     ✓ dplyr 1.0.9
## ✓ tidyr 1.2.0      ✓ stringr 1.4.0
## ✓ readr 2.1.2      ✓ forcats 0.5.1
```

```
## — Conflicts — tidyverse_conflicts() —
## X dplyr::filter() masks stats::filter()
## X dplyr::lag()     masks stats::lag()
```

```
penguins %>%
  group_by(species) %>%
  summarise(across(where(is.numeric), mean, na.rm = T))
```

| species
<fct> | bill_length_mm
<dbl> | bill_depth_mm
<dbl> | flipper_length_mm
<dbl> | body_mass_g
<dbl> | year
<dbl> |
|------------------|-------------------------|------------------------|----------------------------|----------------------|---------------|
| Adelie | 38.79139 | 18.34636 | 189.9536 | 3700.662 | 2008.013 |
| Chinstrap | 48.83382 | 18.42059 | 195.8235 | 3733.088 | 2007.971 |
| Gentoo | 47.50488 | 14.98211 | 217.1870 | 5076.016 | 2008.081 |

3 rows



10. 분석보고서 작성

- 마크다운 문서 내부에 표와 그림 등을 포함할 수 있음

표와 그림 포함하기

마크다운 문서 안에 ****표****도 포함할 수 있습니다.

```
```{r tables}
library(stargazer)
stargazer(data.frame(penguins), type="text")
```
```



10. 분석보고서 작성

표와 그림 포함하기

마크다운 문서 안에 표도 포함할 수 있습니다.

```
library(stargazer)
```

```
##  
## Please cite as:
```

```
## Hlavac, Marek (2022). stargazer: Well-Formatted Regression and Summary Statistics Tables.
```

```
## R package version 5.2.3. https://CRAN.R-project.org/package=stargazer
```

```
stargazer(data.frame(penguins), type="text")
```

```
##  
## =====  
## Statistic      N      Mean    St. Dev.  Min    Max  
## -----  
## bill_length_mm 342   43.922    5.460   32.100  59.600  
## bill_depth_mm  342   17.151    1.975   13.100  21.500  
## flipper_length_mm 342  200.915   14.062    172    231  
## body_mass_g    342  4,201.754  801.955   2,700   6,300  
## year           344  2,008.029   0.818   2,007   2,009  
## -----
```



10. 분석보고서 작성

마크다운 문서 안에 ****그림****도 포함할 수 있습니다.

```
```{r figures, echo=FALSE}
ggplot(data = penguins,
 mapping = aes(x = flipper_length_mm, y = body_mass_g)) +
 geom_point(aes(color = species, shape = species), size = 3, alpha = 0.8) +
 scale_color_manual(values = c("darkorange", "purple", "cyan4")) +
 labs(title = "Penguin size, Palmer Station LTER",
 subtitle = "Flipper length and body mass for Adelie, Chinstrap and Gentoo Penguins",
 x = "Flipper length (mm)",
 y = "Body mass (g)",
 color = "Penguin species",
 shape = "Penguin species") +
 theme_minimal()
```
```

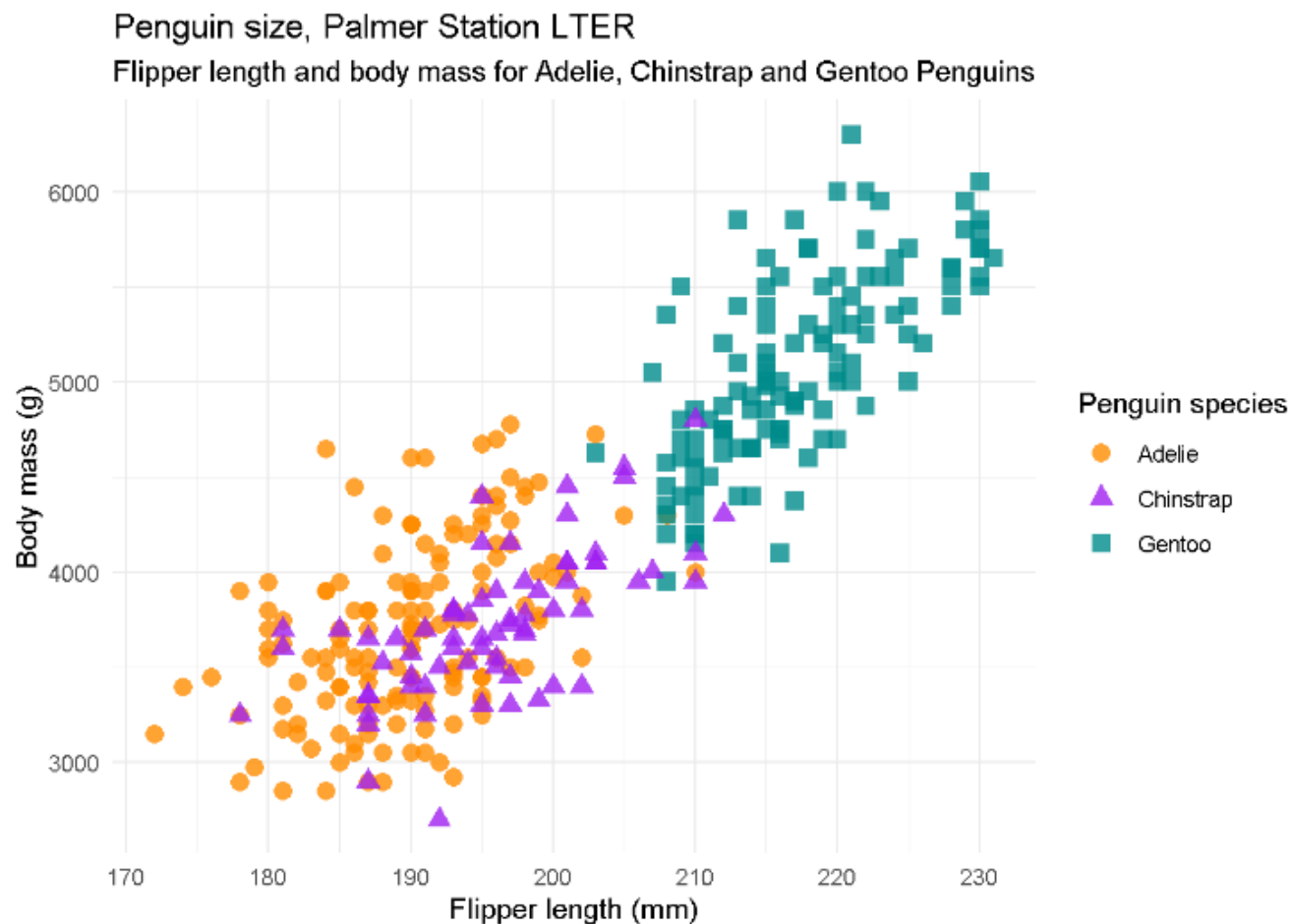
Note that the ``echo = FALSE`` parameter was added to the code chunk to prevent printing of the R code that generated the plot.



10. 분석보고서 작성

마크다운 문서 안에 그림도 포함할 수 있습니다.

```
## Warning: Removed 2 rows containing missing values (geom_point).
```

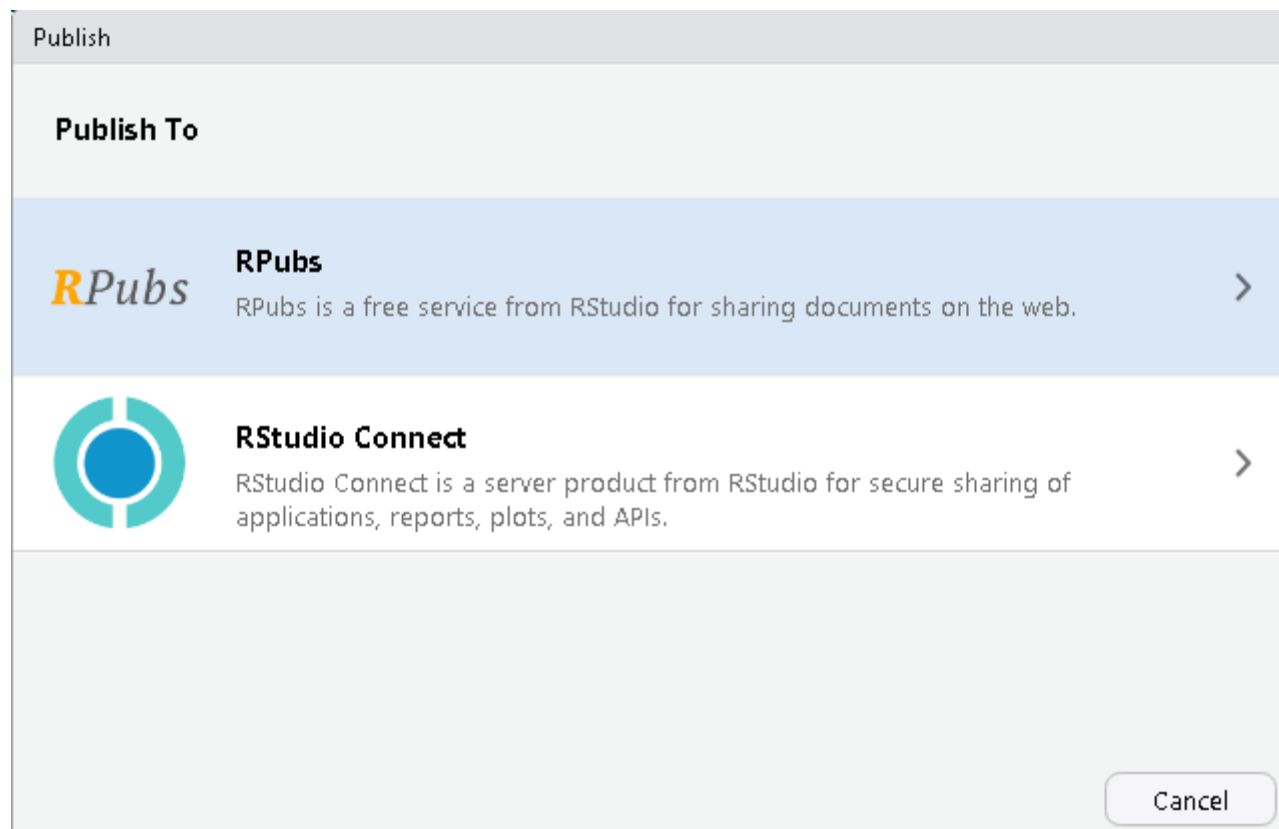


Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.



10. 분석보고서 작성

- 우측 상단의 Publish 버튼으로 RPubS, RStudio Connect 등에 R 마크다운 문서를 배포할 수 있음





■ R Markdown Cheat Sheets:

- [Help > Cheat Sheets > R Markdown Cheat Sheet](#)

rmarkdown

What is markdown?

• Rmd files: Develop your code and ideas side-by-side in a single document. Run code as individual chunks or as an entire document.

• Dynamic Documents: Knit together reports, tables, and results with narrative text. Render to a variety of formats like HTML, PDF, MS Word, or MS PowerPoint.

• Reproducible Research: Upload, link to, or attach your report to share. Anyone can read or run your code to reproduce your work.

SOURCE EDITOR

RENDERED OUTPUT

Workflow

1. Open a new Rmd file in the RStudio IDE by going to **File > New File > R Markdown...**
2. Embed code in chunks. Run code by chunk, or all at once.
3. Write text and add tables, figures, images, and citations. Format with Markdown syntax or the RStudio Visual Markdown Editor.
4. Set output format(s) and options in the YAML header. Customize themes or add parameters to execute or add interactivity with Shiny.
5. Save and render the whole document. Knit periodically to preview your work as you write.
6. Share your work!

VISUAL EDITOR

Write with Markdown

The syntax on the left renders as the output on the right.

Plain text.
 End a line with two spaces to start a new paragraph.
 Also end with a backslash to make a new line.
 "italics" and "bold"
 superscript²/subscript₂
 strikethrough
 escaped: "\\\\"
 endash: " -- " endash: " --- "

Header 1
Header 2

Header 6

1. ordered list
 2. Item 2
 3. Item 2a (indent 1 tab)
 4. Item 2b (indent 2 tab)

1. ordered list
 2. Item 2
 3. Item 2a (indent 1 tab)
 4. Item 2b (indent 2 tab)

<http://www.rstudio.com/>
 This is a link.

Caption.
 verbatim code
 multiple lines of verbatim code
 > block quotes

equation: $x^2 + 1 = 0$
 equation block:

$$E = mc^2$$

 horizontal rule:

HTML Tables
 # Results [source]
 # F003 test
 test
 # Tables
 more text

Results
 Plots Tables
 text

Embed Code with knitr

CODE CHUNKS

Surround code with ````{r}` or use the Insert Code Chunk button. Add a chunk label and/or chunk options like `eval=FALSE` or `cache=TRUE`.

```
```{r chunk-label, include=FALSE}
summary(mtcars)
```
```

SET GLOBAL OPTIONS

Set options for the entire document in the first chunk.

```
```{r include=FALSE}
knitr::opts_chunk$set(message = FALSE)
```
```

INLINE CODE

Insert `r`code`` into text sections. Code is evaluated at render and results are saved as text.

````{r}`getVersion()```` → ````{r}`getVersion()````

## R Studio

RStudio® is a trademark of RStudio, PBC • CC BY SA RStudio • info@rstudio.com • 844-448-1212 • rstudio.com • Learn more at [markdowndownload.com](http://markdowndownload.com) • markdowndownload.com • Updated: 2021-01-01

*Any Questions?*

