

R03. R Markdown (basics)

Getting started

Sim, Min Kyu, Ph.D.
mksim@seoultech.ac.kr



서울과학기술대학교 데이터사이언스학과

1 Getting started

2 Formats

3 .Rmd 파일의 구성과 문법

4 Cheatsheets

Getting started

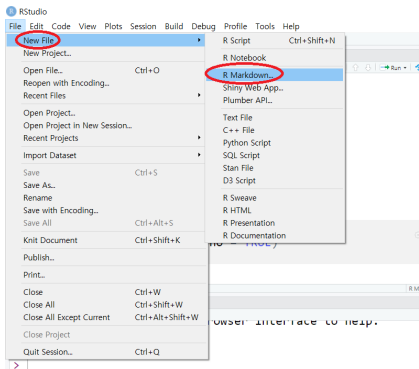
Motivation

- 분석 결과물은 공유되어야 하기에 문서화는 필수적인 과정이다.
- 분석 → 결과정리 → 문서화의 과정을 하나로 연결시킬 수는 없을까?
- Workflow
 - Old: 계산 → 표, 그림 정리 → 워드프로세서에 정리 → 다시 계산 → ...
 - New: rmarkdown
- Tools의 관점
 - Old: 에디터, 콘솔, 아웃풋 윈도우, 엑셀, 워드프로세서
 - New: rmarkdown in R Studio

Getting Started (.html)

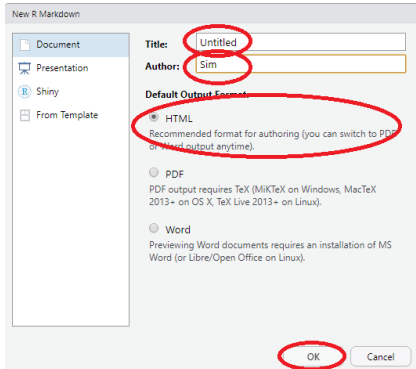
Step 1

File → New File → R Markdown



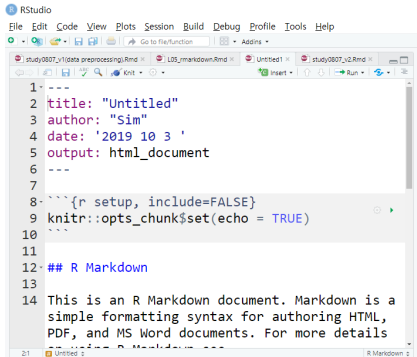
Step 2

Document → Title & Author → html



Step 3

템플릿 소스 코드 등장

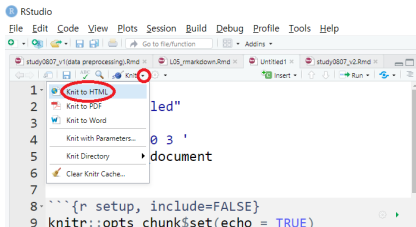


The screenshot shows the RStudio editor with a new R Markdown document. The code is as follows:

```
1 ---
2 title: "Untitled"
3 author: "Sim"
4 date: '2019 10 3 '
5 output: html_document
6 ---
7
8 ```{r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = TRUE)
10 ```
11
12 ## R Markdown
13
14 This is an R Markdown document. Markdown is a
15 simple formatting syntax for authoring HTML,
16 PDF, and MS Word documents. For more details
17 see the R Markdown cheatsheet.
```

Step 4

Knit → Knit to html

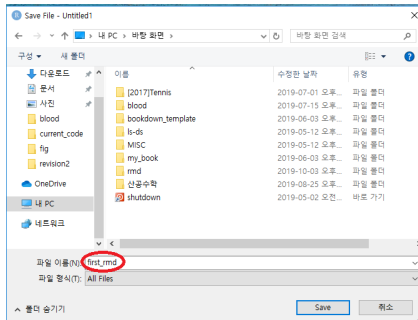


The screenshot shows the RStudio interface with the 'Knit' menu open. The 'Knit to HTML' option is highlighted with a red circle. The code in the background is the same as in Step 3.

```
1 Knit to HTML
2 Knit to PDF
3 Knit to Word
4 Knit with Parameters...
5 Knit Directory
6 Clear Knitr Cache...
7
8 ```{r setup, include=FALSE}
9 knitr::opts_chunk$set(echo = TRUE)
```

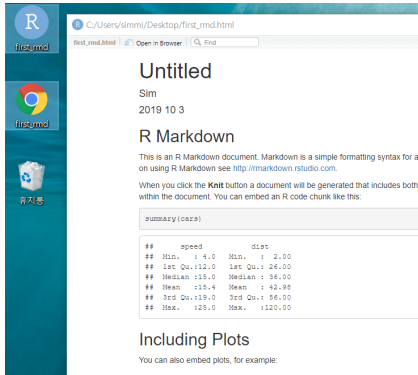
Step 5

파일 이름 지정



Step 6

rmd 파일과 html 파일 생성됨



.Rmd VS html

The image shows the RStudio Source Editor on the left and a web browser on the right, displaying the rendered HTML output of an R Markdown document. Red boxes and arrows highlight key features and syntax elements.

Source Editor (Left):

- Line 1: `---`
- Line 2: `title: "Untitled"` (Red box, arrow pointing to "1. 제목, 저자, 날짜")
- Line 3: `author: "Sim"`
- Line 4: `date: '2019 10 3 '`
- Line 5: `output: html_document`
- Line 6: `---`
- Line 8: `{r setup, include=FALSE}`
- Line 9: `knitr::opts_chunk$set(echo = TRUE)`
- Line 10: `---`
- Line 12: `## R Markdown` (Red box, arrow pointing to "2. 섹션의 제목을 정할 때에는 #기호를 사용. 두번을 쓴 경우를 'level 2 헤더'라고 함")
- Line 14: `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`
- Line 15: `<http://rmarkdown.rstudio.com>` (Red box, arrow pointing to "3. <>기호를 이용해 하이퍼링크 생성")
- Line 16: `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:` (Red box, arrow pointing to "4. **마크다운 굵은 글씨**")
- Line 18: `{r cars}`
- Line 19: `summary(cars)` (Red box, arrow pointing to "5. R chunk 부분 cars는 해당 r chunk의 이름. 코드의 결과가 같이 문서에 등장함. R chunk를 시작할 때에는 백틱 (`) 3회와 종괄호 만에 r을 넣고, 종료할 때에는 백틱 (`)를 다시 3회 적어준다")
- Line 20: `---`

Web Browser (Right):

- Title: Untitled
- URL: C:/Users/simmi/...
- Content: Untitled, Sim, 2019 10 3
- Section: R Markdown
- Text: 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>.
- Text: 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:
- Code chunk: `summary(cars)`
- Table output:

	speed	dist
## Min.	4.0	Min. : 2.00
## 1st Qu.	12.0	1st Qu.: 26.00
## Median	15.0	Median : 36.00
## Mean	15.4	Mean : 42.98
## 3rd Qu.	19.0	3rd Qu.: 56.00

RStudio Source Editor

```

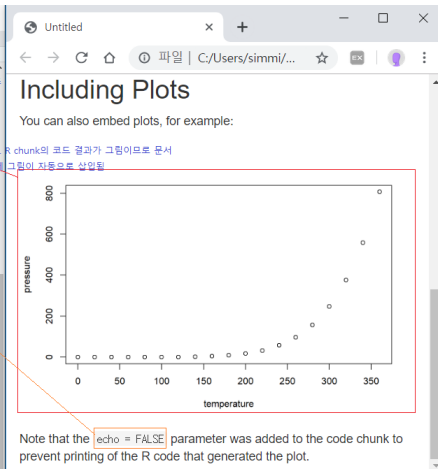
21
22 ## Including Plots
23
24 You can also embed plots, for example:
25
26 ```{r pressure, echo=FALSE}
27 plot(pressure)
28 ```
29
30 Note that the `echo = FALSE` parameter was
31 added to the code chunk to prevent printing of
32 the R code that generated the plot.
33
34
35
36
37
38
39
40
41
42
43
44
45

```

6. r chunk의 시작에 echo=FALSE를 넣어주면 문서에 코드가 보이지 않음

7. R chunk의 코드 결과가 그림이므로 문서에 그림이 자동으로 삽입됨

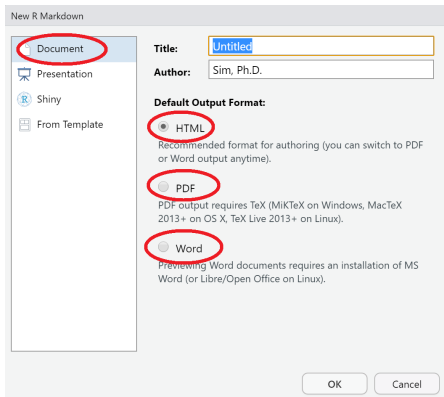
8. 백틱 (')으로 감싸준 글자는 typesetting font (monospace)로 바뀌어서 코드임을 구별할 수 있게 해줌.



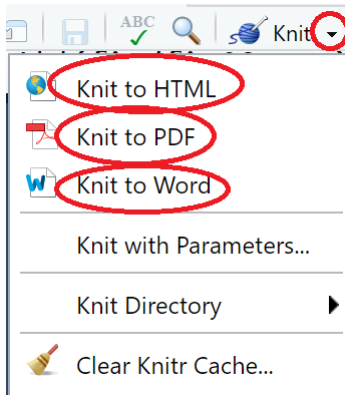
Formats

Documents

- 다양한 포맷의 문서를 생성할 수 있음

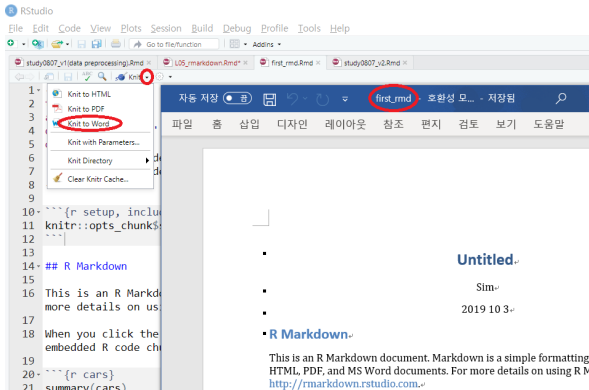


- 기존의 문서에서 버튼을 누르는 것만으로도 포맷을 변경할 수 있음



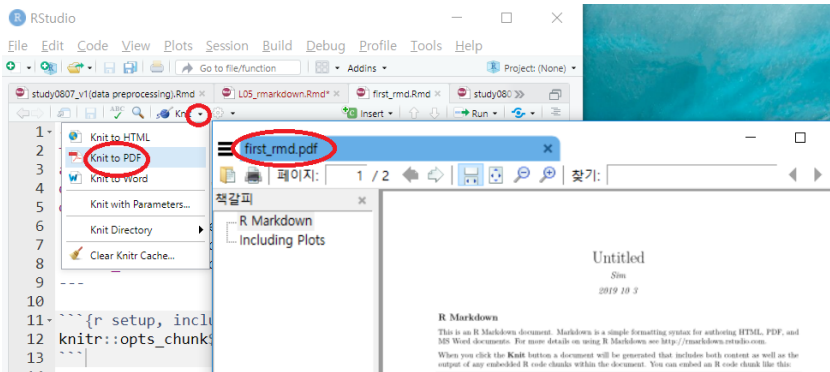
.docx

- MS Word 프로그램이 설치되어 있지 않더라도 .docx 생성 가능



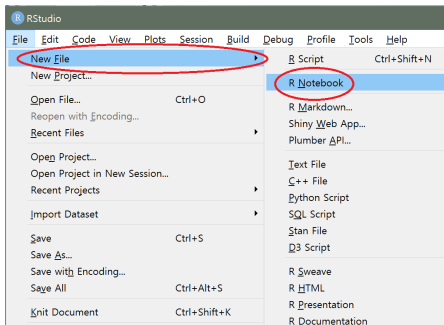
.pdf

- pdf 문서형식은 편집이 불가능하며 가장 professional한 document format임
- 조판프로그램인 Miktek(영문 only), texlive(한글 지원)등의 프로그램의 설치가 필요

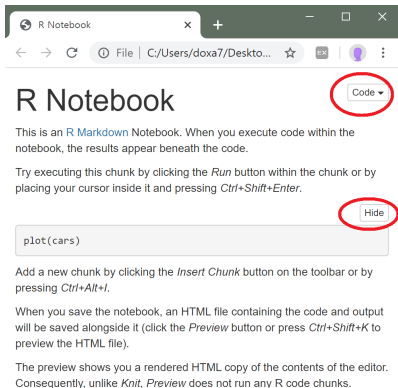


R Notebook

- .html과 유사한 방식으로 웹문서 생성

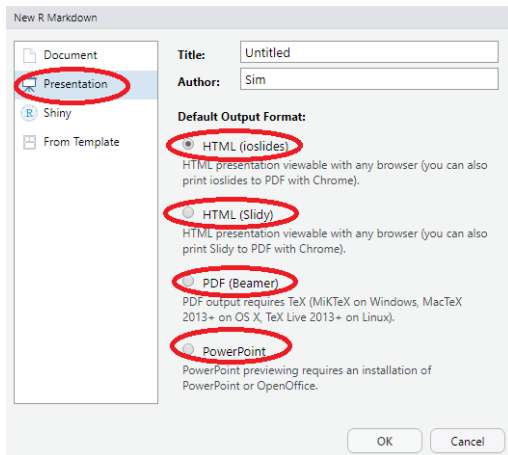


- Code 버튼에서 .Rmd파일을 다운로드 받을 수도 있고
- 독자가 코드를 Show/Hide하면서 문서를 읽을 수 있음



Slides

- 다양한 포맷의 슬라이드를 생성할 수 있음

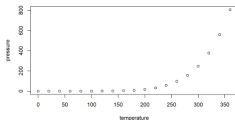


ioslides (.html)

Slide with Bullets

- Bullet 1
- Bullet 2
- Bullet 3

Slide with Plot



Untitled

Sim
2019 10 3

3/5

5/5

R Markdown

This is an R Markdown presentation. 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.

2/5

Slide with R Output

```
summary(cars)
```

```
##      speed          dist  
##  Min.   : 4.0      Min.   : 2.00  
## 1st Qu.:12.0      1st Qu.: 26.00  
##   Median :15.0      Median : 36.00  
##    Mean  :15.4      Mean   : 42.98  
## 3rd Qu.:19.0      3rd Qu.: 56.00  
##   Max.  :25.0      Max.   :120.00
```

4/5

Slidy (.html)

Untitled

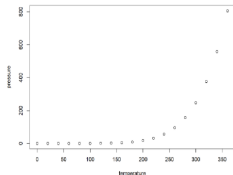
Sim

2019 10 3

Slide with Bullets

- Bullet 1
- Bullet 2
- Bullet 3

Slide with Plot



R Markdown

This is an R Markdown presentation. 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.

Slide with R Output

```
summary(mtcars)
##      speed          dist
##  Min.   : 4.0    Min.   : 2.00
##  1st Qu.:12.0    1st Qu.: 26.00
##  Median :15.0    Median : 36.00
##  Mean   :15.4    Mean   : 42.98
##  3rd Qu.:19.0    3rd Qu.: 58.00
##  Max.   :25.0    Max.   :120.00
```

beamer (.pdf)

Unrolled

Kim

2019-10-13

Slide with Bullets

- Bullet 1
- Bullet 2
- Bullet 3

Slide with Plot



R Markdown

This is an R Markdown presentation. 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>.

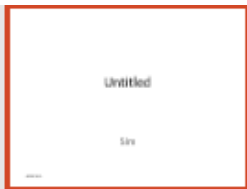
When you click the **Knit** button a document will be generated that includes both content as well as the output of any embedded R code chunks within the document.

Slide with R Output

```
##> summary()
##      speed      dist
##  Min.   : 4.2   Min.   : 2.00
##  1st Qu.:14.4   1st Qu.: 16.00
##  Median :23.2   Median : 26.00
##  Mean   :31.4   Mean   : 42.98
##  3rd Qu.:39.2   3rd Qu.: 58.00
##  Max.   :45.4   Max.   :148.00
```

PowerPoint (.pptx)

1



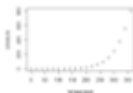
3

Slide with Bullets

- Bullet 1
- Bullet 2
- Bullet 3

5

Slide with Plot



2

R Markdown

This is an R Markdown presentation. 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.

4

Slide with R Output

```
##> summary()
##> log2(x)
##> log2(x) <- 2:5      log2(x) <- 2:5
##> log2(x) <- 2:5      log2(x) <- 2:5
##> log2(x) <- 2:5      log2(x) <- 2:5
##> log2(x) <- 2:5      log2(x) <- 2:5
##> log2(x) <- 2:5      log2(x) <- 2:5
```

.Rmd 파일의 구성과 문법

3가지 구성 요소

1. YAML Header: Rmd파일의 처음에 등장하며, 문서의 포맷을 정하는 기능
2. R Chunk: R언어로 작성하며 배경 색이 다름
3. MD (Markdown) Chunk: 사람의 언어로 입력

1. Yaml Header

- Rmd파일의 처음에 등장하며, ---로 둘러싸인 영역
- 문서의 title, author, date를 지정
- 파일 포맷 및 옵션 지정 가능
- 하위 항목은 빈칸 두 번으로 들여쓰기

Documents

```
1 ---
2 title: "Untitled"
3 author: "Sim"
4 date: '2019 10 3 '
5 output:
6   pdf_document: default
7   html_document: default
8   word_document: default
9 ---
```

Slides

```
1 ---
2 title: "Untitled"
3 author: "Sim"
4 date: '2019 10 3 '
5 output:
6   beamer_presentation: default
7   ioslides_presentation: default
8   slidy_presentation: default
9 ---
```

pptx

```
1 ---
2 title: "Untitled"
3 author: "Sim"
4 date: '2019 10 3 '
5 output: powerpoint_presentation
6 ---
```

2. R Chunk

- R언어로 작성하며 Color로 구분됨
- R Chunk를 삽입하는 방법
 1. 백틱 (키보드의 “1” 키 왼쪽에 위치함)을 3번 입력하고 중괄호 안에 r을 넣어주면 R Chunk를 시작
 2. R Chunk를 종료하기 위해서는 백틱을 3번 입력
 3. 단축키 Ctrl+Alt+I를 사용하면 R Chunk가 삽입됨

● R Chunk의 argument

- 아래 예제의 `pressure`는 해당 R Chunk의 고유한 이름이며, 지정하지 않으면 자동으로 지정됨
- `echo=FALSE`는 코드를 보이지 않게 한다는 의미이며, 다른 여러 종류의 chunk option이 존재함
- Chunk option을 .Rmd 파일 전체에 대해서 global하게 지정할 수도 있음

```
## 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

● R Chunk의 오른쪽 세개의 버튼

1. Setting: Chunk option을 지정
2. Run all chunks above: 위에 있는 모든 R Chunk를 실행
3. Run this chunk: 현재 R Chunk를 실행. Ctrl+Enter 단축키와 같은 기능

● R Chunk options

- `echo=FALSE`를 하면 코드가 문서에 보이지 않음
- `eval=FALSE`를 하면 코드를 실제로는 실행하지 않음
- `results = 'hide'`를 하면 코드의 output을 출력하지 않음

3. MD(markdown) Chunk

- 사람의 언어를 입력해서 데이터 분석의 스토리를 넣어줄 수 있음
- Inline R Code를 사용할 수 있음
 - 백틱으로 감싼후에 r을 넣고 R명령어를 사용하면 됨
 - 결과를 읽고 입력하는 과정을 생략할 수 있기에 빠르게 보고서를 완성할 수 있음
 - R Markdown을 개발하는 motivation이 된 기능

The screenshot shows the RStudio interface with a script editor on the left and a viewer window on the right. The script editor contains R Markdown code, and the viewer shows the rendered HTML output. Red and blue circles and arrows highlight specific elements.

Script Editor (Left):

```
15 |
16 | When you click the **Knit** button a document will be
17 | generated that includes both content as well as the output of
18 | any embedded R code chunks within the document. You can embed
19 | an R code chunk like this:
20 |
21 | ```{r cars}
22 | summary(cars)
23 | ```
```

Viewer Window (Right):

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:

```
summary(cars)
```

| | speed | dist |
|-----------------|-------|----------------|
| ## Min. : | 4.0 | Min. : 2.00 |
| ## 1st Qu.:12.0 | | 1st Qu.: 26.00 |
| ## Median :15.0 | | Median : 36.00 |
| ## Mean :15.4 | | Mean : 42.98 |
| ## 3rd Qu.:19.0 | | 3rd Qu.: 56.00 |
| ## Max. :25.0 | | Max. :120.00 |

• 조사된 차종의 평균 속력은 15.4이다.

- bold와 italic의 사용예시 확인

● Markdown

- plain-text-formatting syntax로 읽고 쓰기 쉬움
- Github등에서 Readme 문서에서도 사용되는 포맷

3. 마크다운 다음으로, 일반 텍스트로 보고서를 작성한다. 마크다운 구문을 사용해서 최종 보고서에 적용할 텍스트 서식을 기술한다.

입력 구문

일반 텍스트

새로운 단락을 시작하려면 줄 마지막을 공백 2개로 끝낸다.

기울인 글씨 and _기울인 글씨_

굵은 글씨 and __굵은 글씨__

윗첨자²

~~취소선~~

[링크](www.rstudio.com)

제목 1

제목 2

제목 3

제목 4

제목 5

제목 6

출력 결과

일반 텍스트

새로운 단락을 시작하려면 줄 마지막을 공백 2개로 끝낸다.

기울인 글씨 and 기울인 글씨

굵은 글씨 and 굵은 글씨

윗첨자²

~~취소선~~

[링크](#)

제목 1

제목 2

제목 3

제목 4

제목 5

제목 6

N자 크기 대신 부호: --

m자 크기 대신 부호: ---

생략: ...

즉시 처리하는 수식: $A = \pi r^2$

이미지:

수평선 (혹은 슬라이드 멈춤):

> 인용 블록

* 순서없는 목록

* 항목 2

+ 하위 항목 1

+ 하위 항목 2

1. 순서있는 목록

2. 항목 2

+ 하위 항목 1

+ 하위 항목 2

표 제목 | 두번째 제목

----- | -----

표 칸 | 칸 2

칸 3 | 칸 4

N자 크기 대신 부호: -

m자 크기 대신 부호: ---

생략: ...

즉시 처리하는 수식: $A = \pi r^2$

이미지: 

수평선 (혹은 슬라이드 멈춤):

인용 블록

• 순서없는 목록

• 항목 2

• 하위 항목 1

• 하위 항목 2

1. 순서있는 목록

2. 항목 2

• 하위 항목 1

• 하위 항목 2

표 제목

두번째 제목

표 칸

칸 2

칸 3

칸 4

이미지 삽입

1. R chunk에서 삽입

- Rmd파일이 속한 폴더의 상대경로로 이미지파일 지정

```
```{r, echo=FALSE, fig.align='center', out.height="80%", out.width="70%"}  
knitr::include_graphics("fig/fig07.png")
```
```

2. Markdown chunk에서 삽입

- ![그림제목](fig.png){ width=400px, height=300px }
 - width와 height를 pixel 단위로 지정
 - 띄어쓰기 주의
- ![그림제목](folder/fig.png){ width=40%, height=50% }
 - width와 height를 문서 1페이지에 대한 상대적인 크기로 지정
 - Rmd파일이 속한 폴더의 상대경로로 지정

Cheatsheets

5. 코드내장하기 knitr 구문을 사용하여 R 코드를 보고서에 내장한다.
R이 코드를 실행하고, 보고서를 렌더링할 때 결과를 포함시킨다.

인라인 코드
r 코드를 백틱(`)으로 감싸면.
R이 인라인 코드를 실행한 결과로 대체한다.

2 더하기 2는 'r 2'와 같다.

코드 덩어리
R코드 덩어리를 ```{r}``` 으로 시작하고 ```{}``` 으로 마무리한다.

실행 결과는 다음과 같다.

Here's some code
`dim(iris)`
[1] 150 5

화면 출력 선택 옵션

knitr 선택 옵션을 사용해서 코드 덩어리 출력 스타일을 적용한다.
코드 상단 줄에 내부에 선택 옵션을 저장한다.

Here's some code
```{r eval=FALSE}``  
dim(iris)`  
...

Here's some code  
`dim(iris)`

Here's some code  
```{r echo=FALSE}``  
dim(iris)`
...

Here's some code
[1] 150 5

선택 옵션 기본 설정 효과

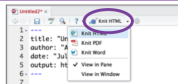
| | | |
|------------|----------------------------------|---------------------------------|
| eval | TRUE | 코드를 평가하고 실행결과를 포함한다. |
| echo | TRUE | 실행결과와 함께 코드를 출력한다. |
| warning | TRUE | 경고메시지를 출력한다. |
| error | FALSE | 오류메시지를 출력한다. |
| message | TRUE | 메시지를 출력한다. |
| tidy | FALSE | 깔끔한 방식으로 코드 형태를 변형한다. |
| results | "markup", "asis", "hold", "hide" | |
| cache | FALSE | 결과값을 캐시해서 향후 실행시 건너뛰게 설정한다. |
| comment | "##" | 주석문자로 출력결과에 서두를 붙인다. |
| fig.width | 7 | 덩어리로 생성되는 그래프에 대한 폭을 인치로 지정한다. |
| fig.height | 7 | 덩어리로 생성되는 그래프에 대한 높이를 인치로 지정한다. |

보다 자세한 사항은 웹사이트를 참조: yihui.name/knitr/

6. 렌더링 최종보고서를 생성하는데 .Rmd 파일을 사용하여 청사진을 제작한다.

두가지 방식으로 보고서를 렌더링한다.

1. `rmarkdown::render("<파일 경로>")` 명령어를 실행한다.
2. RStudio 스크립트 작성창 상단에 **knit HTML** 버튼을 클릭한다.



렌더링 명령을 실행시키면, R은 다음을 수행한다

- 내장된 코드 덩어리를 각각 실행시키고, 실행결과를 보고서에 삽입한다.
- 출력 파일형식에 맞춰 신규 보고서를 생성한다.
- 미리보기로 뷰어창에 출력파일을 연다.
- 작업디렉토리에 출력파일을 저장한다.

7. 인터랙티브 문서 작성한 보고서를 3단계를 거쳐 인터랙티브 Shiny 문서변환.

1. **YAML 헤더에 runtime: shiny**를 추가한다.

title: "Line graph"
output: html_document
runtime: shiny

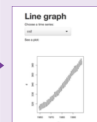
2. 코드 덩어리에, 위젯을 내장하는 **Shiny input** 함수를 추가한다. **Shiny render** 함수를 추가해서 반응형 출력 결과를 내장한다.

```
title: "Line graph"
output: html_document
runtime: shiny

(Choose a time series:
r echo = FALSE)
selectInput("data", "",
  c("co2", "lk"))

See a plot:
r echo = FALSE)
renderPlot({
  d <- get(input$data)
  plot(d)
})
```

3. `rmarkdown::run` 명령어로 렌더링하거나 RStudio **Run Document** 버튼을 클릭한다.



* 주석: 보고서는 Shiny 앱이 된다. 따라서, (인터랙티브 보고서를 위해) `html_document` 혹은 (인터랙티브 발표자료) `ioslides_presentation` 출력형식을 선택한다.

8. Publish 온라인으로 접속하는 사용자와 보고서를 공유한다.

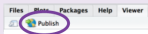
Rpubs.com

RStudio 무료 R 마크다운 게시 사이트를 통해 정적 문서를 공유한다.
www.rpubs.com

ShinyApps.io

Shiny 서버에 인터랙티브 문서를 올려 호스팅한다. 무료와 유료 선택 옵션이 있다.
www.shinyapps.io

RStudio 미리보기 창에 "Publish" 버튼을 클릭하여, `rpubs.com` 사이트에 배포한 후 반별 업데이트로 배포 가능하다.



9. 추가 학습

문서와 예제 - rmarkdown.rstudio.com
추가 기사 - shiny.rstudio.com/articles
@ - blog.rstudio.com
- @rstudio



RStudio® and Shiny™ are trademarks of RStudio, Inc.
© 2015 RStudio, Inc. info@rstudio.com
844-448-1212 rstudio.com

Suggested Readings

- R for Data Science
 - <https://r4ds.had.co.nz>
- [R4DS] Ch21. R Markdown
 - <https://r4ds.had.co.nz/communicate-intro.html>
 - <https://r4ds.had.co.nz/r-markdown.html>

```
"R Markdown basics"
```

```
## [1] "R Markdown basics"
```