A UNIFIED APPROACH FOR WRITING AUTOMATIC REPORTS

Parameterization and Generalization of R-Markdown

Cristina Muschitiello
Niccolò Stamboglis

04_Generalizability

What is a generalizable Rmd?

Generalization definitions from Computer Science:

> *Generalization* is the process of extracting shared characteristics from two or more classes, and combining them into a generalized superclass. Shared characteristics can be attributes, associations, or methods. [1]

> **Generalization**
> the identification, and possible organization, of common properties of abstractions. [2]

CODE ORIENTED SOFTWARE REUSE [3]

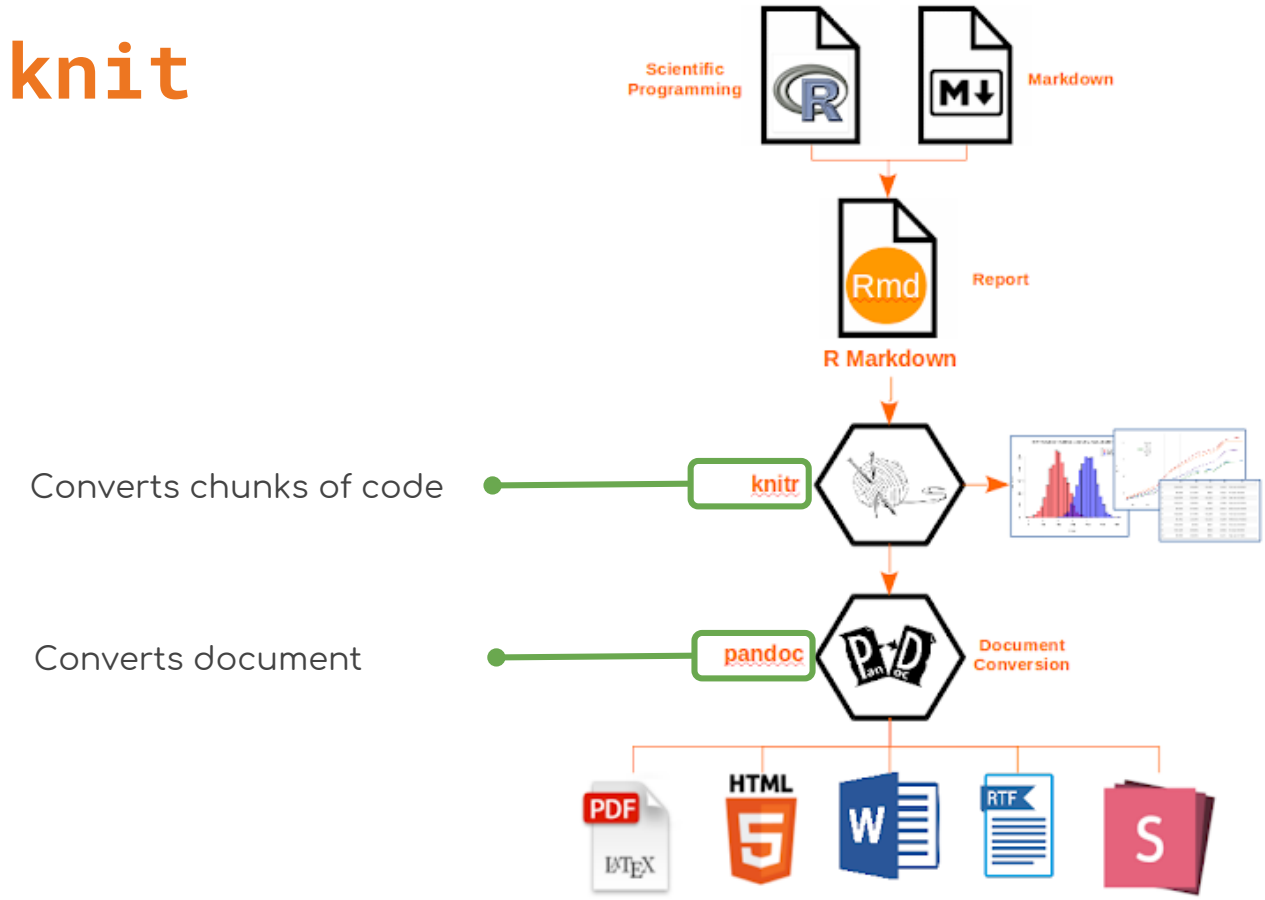
What is a generalizable Rmd?

Generalization -> use the **same Rmd script** to analyse **different facets** of a study by making use of **parameters**, thus **never altering** the Rmd code.

What is a generalizable Rmd?

Generalization -> use the **same Rmd script** to analyse **different facets** of a study by making use of **parameters**, thus **never altering** the Rmd code.

Behind the knit

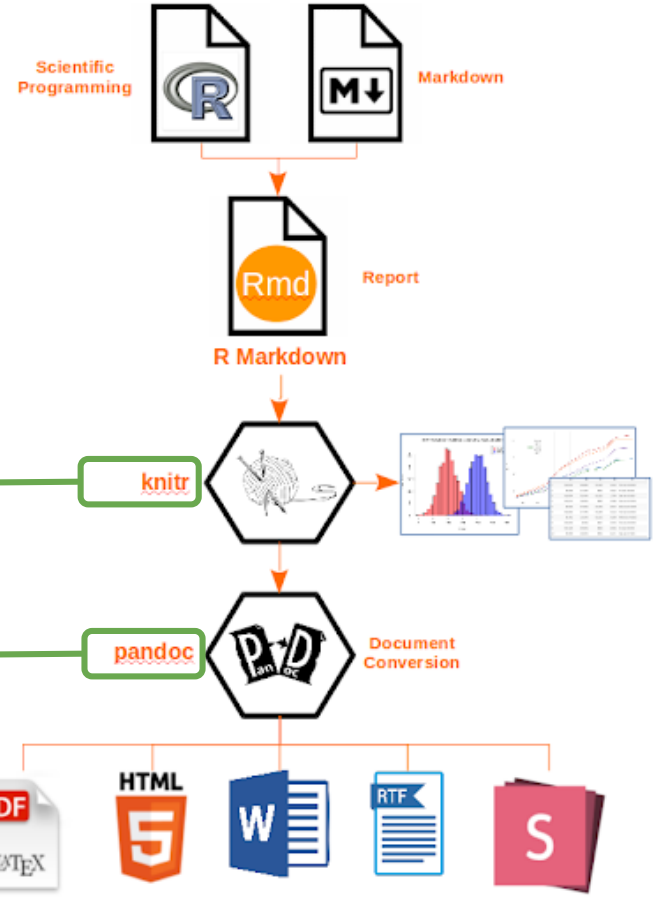


Behind the knit

The **rmarkdown::render()** function

Converts chunks of code

Converts document



rmarkdown.rstudio.com

Render function

Input Rmd/md file
Format (html, pdf)
Output name
Output directory

...

YAML parameters

```
render(
  input,
  output_format = NULL,
  output_file = NULL,
  output_dir = NULL,
  output_options = NULL,
  output_yaml = NULL,
  intermediates_dir = NULL,
  knit_root_dir = NULL,
  runtime = c("auto", "static", "shiny", "shiny_prerendered"),
  clean = TRUE,
  params = NULL,
  knit_meta = NULL,
  envir = parent.frame(),
  run_pandoc = TRUE,
  quiet = FALSE,
  encoding = "UTF-8"
)
```

rmarkdown.rstudio.com

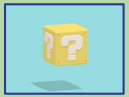
Your turn: Exercise 9

E09_renderRmarkdown.R

The render function run the Rmd contained in the folder named "markdown"

```
rmarkdown::render(
  input="/cloud/project/markdown/GP_report_render_function_s.Rmd",
  output_format = "html_document",
  output_dir = "/cloud/project/solutions",
  params = list(
    month = "07",
    year = 2019,
    gp = "A81005"
  ),
  run_pandoc = TRUE,
  output_file = paste0("/cloud/project/output/GP_report_render_function_s.html")
)
```

- Change output parameters from the render function
- Change output folder



Time: 5 min

! ! GREAT ! !

LET’S MOVE
ANOTHER STEP

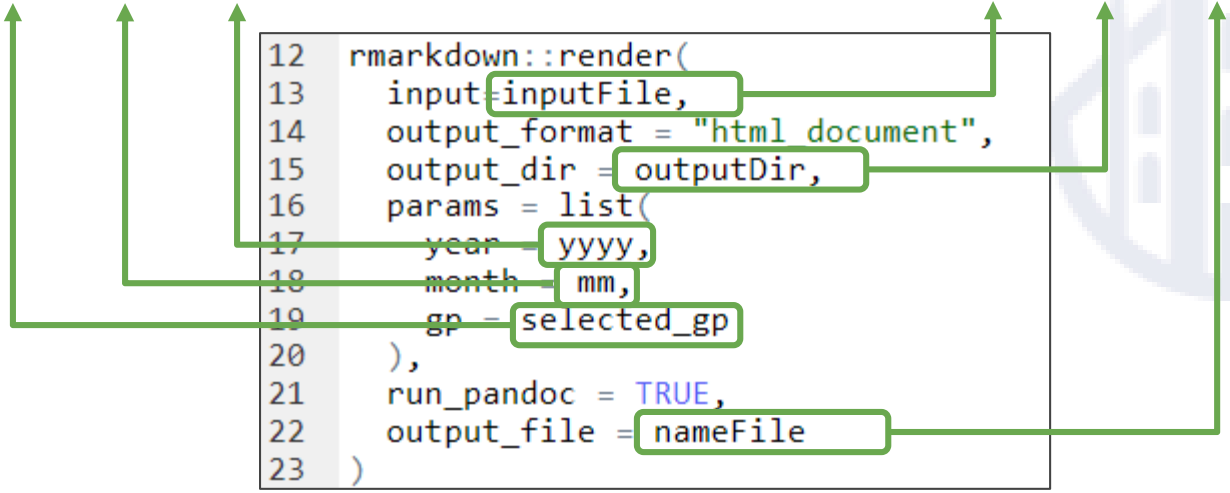
Parameterized render function

DEFINE ALL PARAMETERS OUTSIDE THE RENDER FUNCTION

```
12  rmarkdown::render(
13    input=inputFile,
14    output_format = "html_document",
15    output_dir = outputDir,
16    params = list(
17      year = yyyy,
18      month = mm,
19      gp = selected_gp
20    ),
21    run_pandoc = TRUE,
22    output_file = nameFile
23  )
```

Parameterized render function

```
3  inputFile = "/cloud/project/markdown/GP_report_render_function.Rmd"
4  outputDir = "/cloud/project/output"
5  yyyy = 2019
6  mm = "07"
7  selected_gp = "A81005"
8  nameFile = "GP_report_render_function_E11.html"
```



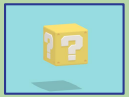
Your turn: Exercise 10

E10_renderRmarkdown_parameterized.R

No need to alter the render function to change output!

```
3  inputFile = "/cloud/project/markdown/GP_report_render_function.Rmd"
4  outputDir = "/cloud/project/output"
5  yyyy = 2019
6  mm = "07"
7  selected_gp = "A81005"
8  nameFile = "GP_report_render_function_E11.html"
```

- Change any parameter



Time: 5 min

! ! SUPER ! !

MOVE TO THE
FINAL LEVEL

Generalized render function

Embed the render function in a function

```
12  rmarkdown::render(
13    input=inputFile,
14    output_format = "html_document",
15    output_dir = outputDir,
16    params = list(
17      year = yyyy,
18      month = mm,
19      gp = selected_gp
20    ),
21    run_pandoc = TRUE,
22    output_file = nameFile
23  )
```

Generalized render function

```
renderReport = function(inputFile, outputDir, yyyy, mm, selected_gp, nameFile){

  rmarkdown::render(
    input=inputFile,
    output_format = "html_document",
    output_dir = outputDir,
    params = list(
      year = yyyy,
      month = mm,
      gp = selected_gp
    ),
    run_pandoc = TRUE,
    output_file = paste0(outputDir,"/",nameFile)
  )
}
```

Generalized render function

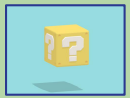
```
renderReport = function(inputFile, outputDir, yyyy, mm, selected_gp, nameFile){

  rmarkdown::render(
    input=inputFile,
    output_format = "html_document",
    output_dir = outputDir,
    params = list(
      year = yyyy,
      month = mm,
      gp = selected_gp
    ),
    run_pandoc = TRUE,
    output_file = paste0(outputDir,"/",nameFile)
  )
}
```

Your turn: Exercise 11

E11_renderFunction.R

```
19
20  # Run the function with chosen parameters
21
22  renderReport(inputFile = "/cloud/project/markdown/GP_report_render_function.Rmd",
23               outputDir = "/cloud/project/output",
24               yyyy = 2019,
25               mm = "07",
26               selected_gp = "A81005",
27               nameFile = "GP_report_render_function_E12.html")
```

- Run the function once and then run it “one shot” with different arguments



Time: 5 min

Ok but… why????

Create parameterized and generalized Rmd skelton reports in any project that can be run for different facets of your analysis without rewriting and/or manually selecting parameters.

Applications

- ➔ Share (reproducible) reports across team
- ➔ Create (reproducible) customer-centered reports
- ➔ Schedule report generation
- ➔ Embed parameterized report in applications
- ➔ ...

Questions

Next up: Prescriptions Project

Source: giphy.com