

Template for contribution to Computo

Example dedicated to R users

Jane Doe <a>0¹ Statistics, Name of Affiliation one John Doe <a>0 Computer Science, Name of Afficiliation two

Date published: 2/22/23 Last modified: 2/22/23

Abstract

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Curabitur posuere vestibulum facilisis. Aenean pretium orci augue, quis lobortis libero accumsan eu. Nam mollis lorem sit amet pellentesque ullamcorper. Curabitur lobortis libero eget malesuada vestibulum. Nam nec nibh massa. Pellentesque porttitor cursus tellus. Mauris urna erat, rhoncus sed faucibus sit amet, venenatis eu ipsum.

Keywords: key1, key2, key3

Contents

| 1 | Intr | Introduction | | | | |
|---|------|---|---|--|--|--|
| | 1.1 | About this document | 2 | | | |
| | 1.2 | Setup a github repository for preparing your submission | 2 | | | |
| | 1.3 | Quarto | 2 | | | |
| | 1.4 | Requirements | 2 | | | |
| | 1.5 | Link with your usual tools | 2 | | | |
| 2 | For | matting | 3 | | | |
| | 2.1 | Basic markdown formatting | 3 | | | |
| | 2.2 | Mathematics | 3 | | | |
| | | 2.2.1 Mathematical formulae | 3 | | | |
| | | 2.2.2 Theorems and other amsthem-like environments | 3 | | | |
| | 2.3 | R Code | 4 | | | |
| | 2.4 | Figures | 4 | | | |
| | 2.5 | Tables | 5 | | | |
| | 2.6 | Handling references | 6 | | | |
| | | 2.6.1 Bibliographic references | 6 | | | |
| | | 2.6.2 Other cross-references | 6 | | | |
| | 2.7 | Advanced formatting | 6 | | | |
| 3 | Fina | alize your submission | 7 | | | |
| | 3.1 | Handle R dependencies with renv | 7 | | | |
| | 3.2 | Continuous integration | 7 | | | |
| | 3.3 | Data and large files | 8 | | | |

¹Corresponding author: janedoe@nowhere.moon

References 8

Session information 8

1 Introduction

1.1 About this document

This document, accompanied with the hopefully finely tuned git repos, provides a template for writing contributions to **Computo** (Computo Team 2020). We show how R code (R Core Team 2020) can be included and how the repository can be set up for triggering github actions for rendering the document, with dependencies handled by renv.

1.2 Setup a github repository for preparing your submission

You can start by clicking the "use this template" button, on the top of the page of the github repository associated to this document. Of course, you can set your repository private during the preparation of your manuscript.

1.3 Quarto

Quarto is a versatile formatting system for authoring documents integrating markdown, LaTeX and code blocks interpreted either via Jupyter or Knitr (thus supporting Python, R and Julia). It relies on the Pandoc document converter.

1.4 Requirements

You need quarto installed on your system and the Computo extension to prepare your document. For the latter, once quarto is installed, run the following to install the extension in the current directory (it creates a _extension directory which is ignored by git thanks to .gitignore by default):

```
quarto add computorg/computo-quarto-extension
```

R and the following R packages must be installed on your computer: knitr, markdown.

1.5 Link with your usual tools

Quarto is expecting a .qmd markdown file, but will also works with a standard Rmarkdown (.Rmd) file. In addition, especially if you are not comfortable with the command line interface, quarto is fully integrated inside the Rstudio IDE so that you can write and build your quarto document inside Rstudio.

Quarto can also process a Jupyter notebook file if you are used to it (it will just require to add the proper YAML metadata²).

Note: More advanced Jupyter-related functionality like Myst/Jupyter book are not supported in this Quarto setup. The markdown syntax inside the Jupyter notebook should follow the Quarto syntax (c.f. below). If you are more comfortable with using Myst/Jupyter book, we provide a specific template but it will requires more formatting work for Computo editorial team, thus highly encourage authors to use the Quarto templates.

²the same metadata as in the template-computo-R. qmd file in the first cell, type "Raw", of the notebook

2 Formatting

This section covers basic formatting guidelines for quarto documents.

To render a document, run quarto render. By default, both PDF and HTML documents are generated:

quarto render template-computo-R.qmd # will render both to html and PDF



Note

To check the syntax of the formatting below, you can use the </> source button at the top left of this document.

2.1 Basic markdown formatting

Bold text or *italic*

- This is a list
- With more elements
- It isn't numbered.

But we can also do a numbered list

- 1. This is my first item
- 2. This is my second item
- 3. This is my third item

2.2 Mathematics

2.2.1 Mathematical formulae

LaTeX code is natively supported³, which makes it possible to use mathematical formulae:

$$f(x_1, ..., x_n; \mu, \sigma^2) = \frac{1}{\sigma \sqrt{2\pi}} \exp\left(-\frac{1}{2\sigma^2} \sum_{i=1}^n (x_i - \mu)^2\right)$$

It is also posible to cross-reference an equation, see Equation 1:

$$D_{x_N} = \frac{1}{2} \begin{bmatrix} x_L^\top & x_N^\top \end{bmatrix} \begin{bmatrix} L_L & B \\ B^\top & L_N \end{bmatrix} \begin{bmatrix} x_L \\ x_N \end{bmatrix}$$

$$= \frac{1}{2} (x_L^\top L_L x_L + 2x_N^\top B^\top x_L + x_N^\top L_N x_N),$$
(1)

2.2.2 Theorems and other amsthem-like environments

Quarto includes a nice support for theorems, with predefined prefix labels for theorems, lemmas, proposition, etc. see this page. Here is a simple example:

³We use lualatex for this purpose.

Theorem 2.1 (Strong law of large numbers). *The sample average converges almost surely to the expected value:*

$$\overline{X}_n \xrightarrow{a.s.} \mu \quad \text{when } n \to \infty.$$

See Theorem 2.1.

2.3 R Code

Quarto uses either Jupyter or knitr to render code chunks. This can be triggered in the yaml header. In this tutorial, we use knitr (R and packages knitr, markdown must be installed on your computer).

```
title: "My Document"
author "Jane Doe"
```

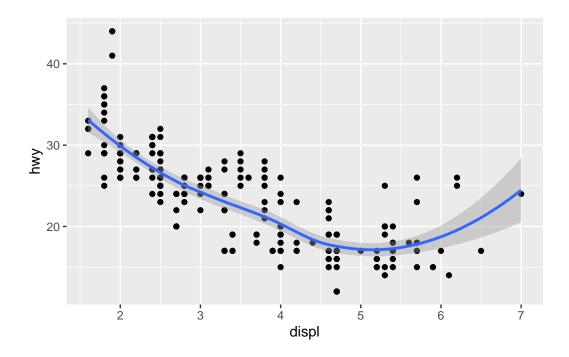
R code (R Core Team 2020) chunks may be embedded as follows:

```
x <- rnorm(10)
```

2.4 Figures

Plots can be generated as follows:

```
library("ggplot2")
p <- ggplot(mpg, aes(displ, hwy)) +
   geom_point() +
   geom_smooth()
p</pre>
```



It is also possible to create figures from static images:

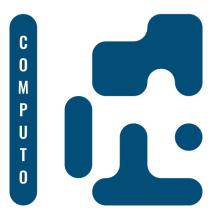


Figure 1: Computo logo (label)

Note: Until Quarto version 1.3+ is released, including a remote image (from a web URL) in a document (like the image above) will work in the rendered HTML document but will generate an error when building the PDF document (c.f. related bug report).

2.5 Tables

Tables (with label: Otbl-mylabel renders Table 1) can be generated with markdown as follows

Table 1: my table caption

| Tables | Are | Cool |
|----------|--------------|--------|
| col 1 is | left-aligned | \$1600 |

| Tables | Are | Cool |
|----------|---------------|------|
| col 2 is | centered | \$12 |
| col 3 is | right-aligned | \$1 |

Table can also be generated by some code, for instance with knitr here:

```
knitr::kable(summary(cars), caption = "Table caption.")
```

Table 2: Table caption.

| 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 |

Handling references 2.6

Bibliographic references

References are displayed as footnotes using BibTeX, e.g. [@computo] will be displayed as (Computo Team 2020), where compute is the bibtex key for this specific entry. The bibliographic information is automatically retrieved from the .bib file specified in the header of this document (here: references.bib).

2.6.2 Other cross-references

As already (partially) seen, Quarto includes a mechanism similar to the bibliographic references for sections, equations, theorems, figures, lists, etc. Have a look at this page.

2.7 **Advanced formatting**

Advanced formatting features are possible and documented (including interactive plots, pseudo-code, (Tikz) diagrams, Lua filters, mixing R + Python in the same document), but are beyond the scope of this simple introduction. We point several entries in this direction.



More information

- The Quarto web site for comprehensive documentation, including:
 - Tutorial
 - User guide
 - Options reference
- The template distributed with the Computo Quarto extension, which uses such advanced features.
- Our mock version of the t-SNE paper, a full and advanced example using Python and

the Jupyter kernel.

• The previously published papers in Computo can be used as references.

3 Finalize your submission

3.1 Handle R dependencies with renv

To make your work reproducible, you need to fix the packages and environment used to run your analysis. For the R system, the renv package is one of the possible reliable method, supported by the community. You basically need a couple of commands to setup your environment on your local machine. First,

```
renv::init()
```

will initialize your repository. Then you just need to install the dependencies required to run your contribution, for instance,

```
renv::install("ggplot2") # or equivalently install.packages("ggplot2")
```

Non-CRAN packages (*e.g.* Github packages) can be used. Once you are done, you can fix everything with the command

```
renv::snapshot()
```

! Important

The only file that needs to be versioned by git is renv.lock. By default, the rest is ignored thanks to .gitignore.

More details for using renv can be found either

- · on the renv packge webpage, or
- · on the quarto page dedicated to environments

3.2 Continuous integration

The repository associated with this template is pre-configure to trigger an action on push that performs the following:

- 1. Check out repository on the ubuntu-latest machine
- 2. Install quarto and dependencies, including the Computo extension
- 3. Install R and dependencies with renv, using your renv.lock file
- 4. Render your .qmd file and Publish the results on a gh-page (both HTML and PDF)

The file .github/workflows/build_n_publish.yml is largely inspired from this file.

Once this is successful, you are ready to submit your manuscript to the Computo submission platform.

Warning

The first time, you possibly need to create the branch for the action to work. This can be done by running the following command from your computer, in your git repository:

```
quarto publish gh-pages
```

Then, set the branch gh-page as the source of your github page, and trigger the action to check that everything works fine.

3.3 Data and large files

If your submission materials contain files larger than 50MB, **especially data files**, they won't fit on a git repository as is. For this reason, we encourage you to put your data or any materials you deem necessary on an external "open data" centered repository hub such a Zenodo or OSF.

References

Computo Team. 2020. "Computo: Reproducible Computational/Algorithmic Contributions in Statistics and Machine Learning."

R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Session information

sessionInfo()

```
R version 4.2.2 (2022-10-31)
Platform: x86_64-pc-linux-gnu (64-bit)
Running under: Ubuntu 22.04.1 LTS
Matrix products: default
        /usr/lib/x86_64-linux-gnu/openblas-pthread/libblas.so.3
LAPACK: /usr/lib/x86_64-linux-gnu/openblas-pthread/libopenblasp-r0.3.20.so
locale:
 [1] LC_CTYPE=C.UTF-8
                            LC_NUMERIC=C
                                                   LC_TIME=C.UTF-8
 [4] LC_COLLATE=C.UTF-8
                           LC_MONETARY=C.UTF-8
                                                  LC_MESSAGES=C.UTF-8
 [7] LC PAPER=C.UTF-8
                           LC NAME=C
                                                   LC ADDRESS=C
                            LC MEASUREMENT=C.UTF-8 LC IDENTIFICATION=C
[10] LC_TELEPHONE=C
attached base packages:
[1] stats
             graphics grDevices datasets utils
                                                      methods
                                                                base
other attached packages:
[1] ggplot2_3.4.0
```

loaded via a namespace (and not attached):

| [1] | knitr_1.42 | magrittr_2.0.3 | splines_4.2.2 | munsell_0.5.0 |
|------|-----------------|-----------------------------|----------------|-----------------|
| [5] | lattice_0.20-45 | <pre>colorspace_2.1-0</pre> | R6_2.5.1 | rlang_1.0.6 |
| [9] | fastmap_1.1.0 | fansi_1.0.4 | tools_4.2.2 | grid_4.2.2 |
| [13] | nlme_3.1-162 | gtable_0.3.1 | mgcv_1.8-41 | xfun_0.37 |
| [17] | utf8_1.2.3 | cli_3.6.0 | withr_2.5.0 | htmltools_0.5.4 |
| [21] | yaml_2.3.7 | digest_0.6.31 | tibble_3.1.8 | lifecycle_1.0.3 |
| [25] | Matrix_1.5-3 | farver_2.1.1 | vctrs_0.5.2 | glue_1.6.2 |
| [29] | evaluate_0.20 | rmarkdown_2.20 | labeling_0.4.2 | compiler_4.2.2 |
| [33] | pillar_1.8.1 | scales_1.2.1 | jsonlite_1.8.4 | renv_0.16.0 |
| [37] | pkgconfig_2.0.3 | | | |