

Introduction

Formating the notebook

Configuring binder

Session information

References

Template for contribution to Computo

Template based on the Rmarkdown system

Computo team

-10

Introduction

What this document

This document provides a Rmarkdown¹ template for contributions to the **Computo** Journal. We show how R and python code can be included.

This document serves as a documentation for configuring the github repository which will host the Rmarkdown source of your manuscript and prove us the reproducibility of your work. To this end, we use binder to generate the rendering of your manuscript both in HTML and PDF.

Computo template

Contributions to Computo require both scientific content (typically equations, codes and figures) and a proof that the content is reproducible. This is achieved via the standard notebook systems available for R, Python and Jupyter notebook and Rmarkdown), coupled with the binder building system.

Your submission is thus a git(hub) repository typically containing

the source of the notebook, which may be a Rmarkdown document (such as the present document) or a Jupyter notebook + Myst **link to other templates when available**;

auxiliary files, e.g.:

- › a BibTeX file (e.g. `./template-computo-Rmarkdown.bib`)

- › some static figures in the `figs/` subdirectory (e.g. `figs/picture.png`)

configuration files for the binder environment to setup the machine that will build the final notebook file (HTML and/or PDF)

This document explains how to:

nat a notebook with Rmarkdown
tionally) configure the binder environment

oiled notebook (pdf file) will be generated directly in the github repository (via a github action), ready
mitted to the Computo submission platform (<https://computo.scholasticahq.com/>).

ce in writting your manuscript

esure that you are able to build your manuscript as a regular notebook on your system. Then you can
figure the binder environment (which we configure to use an Ubuntu machine with latest LTS release).

nating the notebook

on is about writing a notebook with the Rmarkdown system, typically for R users.

rkdown basics

uickly cover the most basic features of Rmarkdown, that is, formatting text with markdown, math
via MathJax and bibliographical references via *BibT_EX*.

wn (<http://rmarkdown.rstudio.com>) is a simple formatting system for authoring HTML and PDF
ts, that relies on the markdown markup language.

the document as HTML within Rstudio, click the **Knit** button. A document will be generated that
both content as well as the output of any embedded R code chunks within the document.
ely, the shortcut Ctrl + Maj + K will produces the same result.

ematical formulae

ode is natively supported, which makes it possible to use mathematical formulae:

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

ences

es are displayed as footnotes using bibtex, e.g. [`@computo`] will display as (Computo Team 2020),
omputo is the bibtex key for this entry. The bibliographic information is automatically retrieved from
file specified in the header of this document (here: `template-computo-Rmarkdown.bib`).

de

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

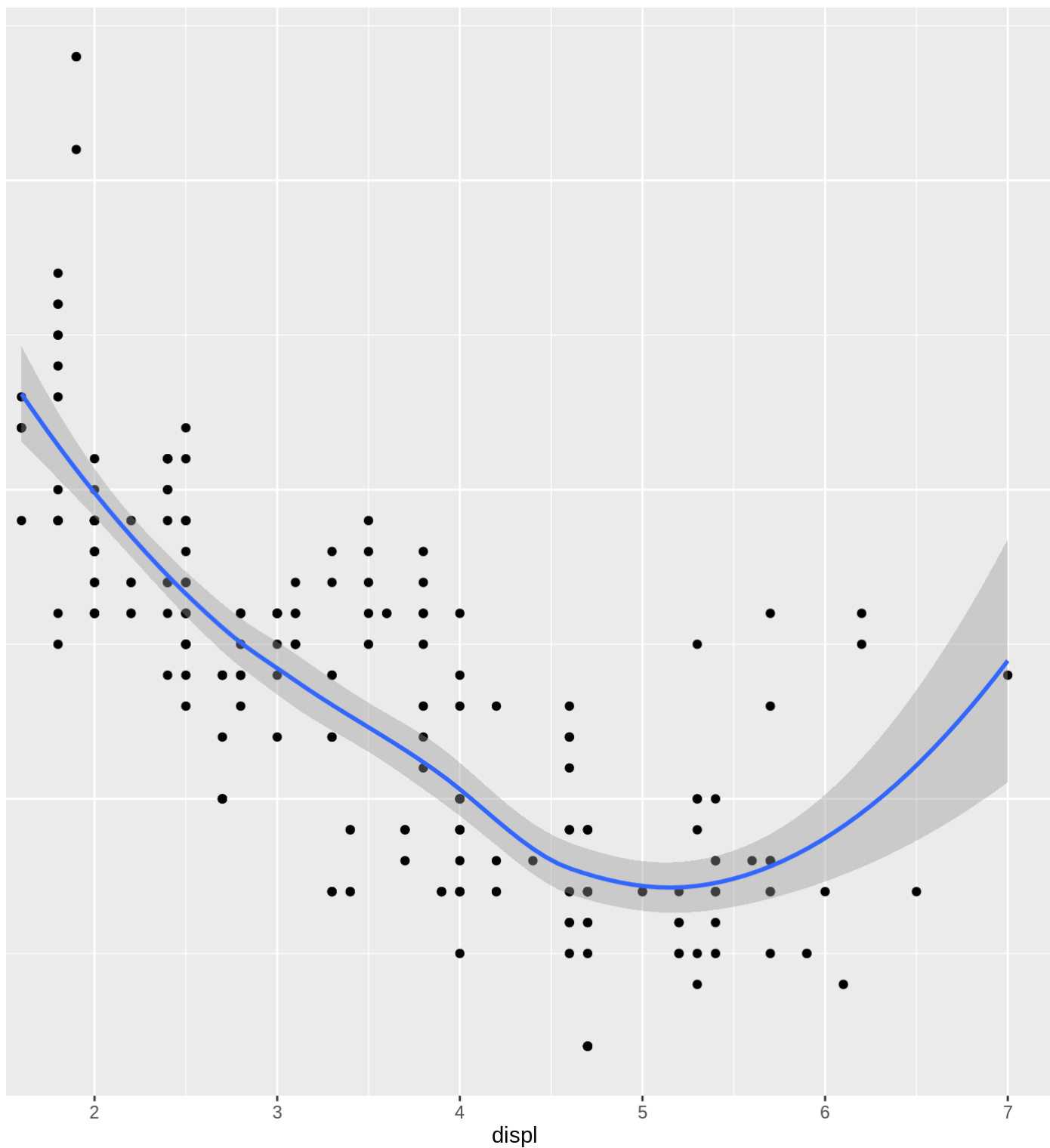
```
library(summary(cars))
```

ed	dist
l. : 4.0	Min. : 2.00
Qu.:12.0	1st Qu.: 26.00
dian :15.0	Median : 36.00
an :15.4	Mean : 42.98
Qu.:19.0	3rd Qu.: 56.00
x. :25.0	Max. :120.00

ing Plots

be generated as:

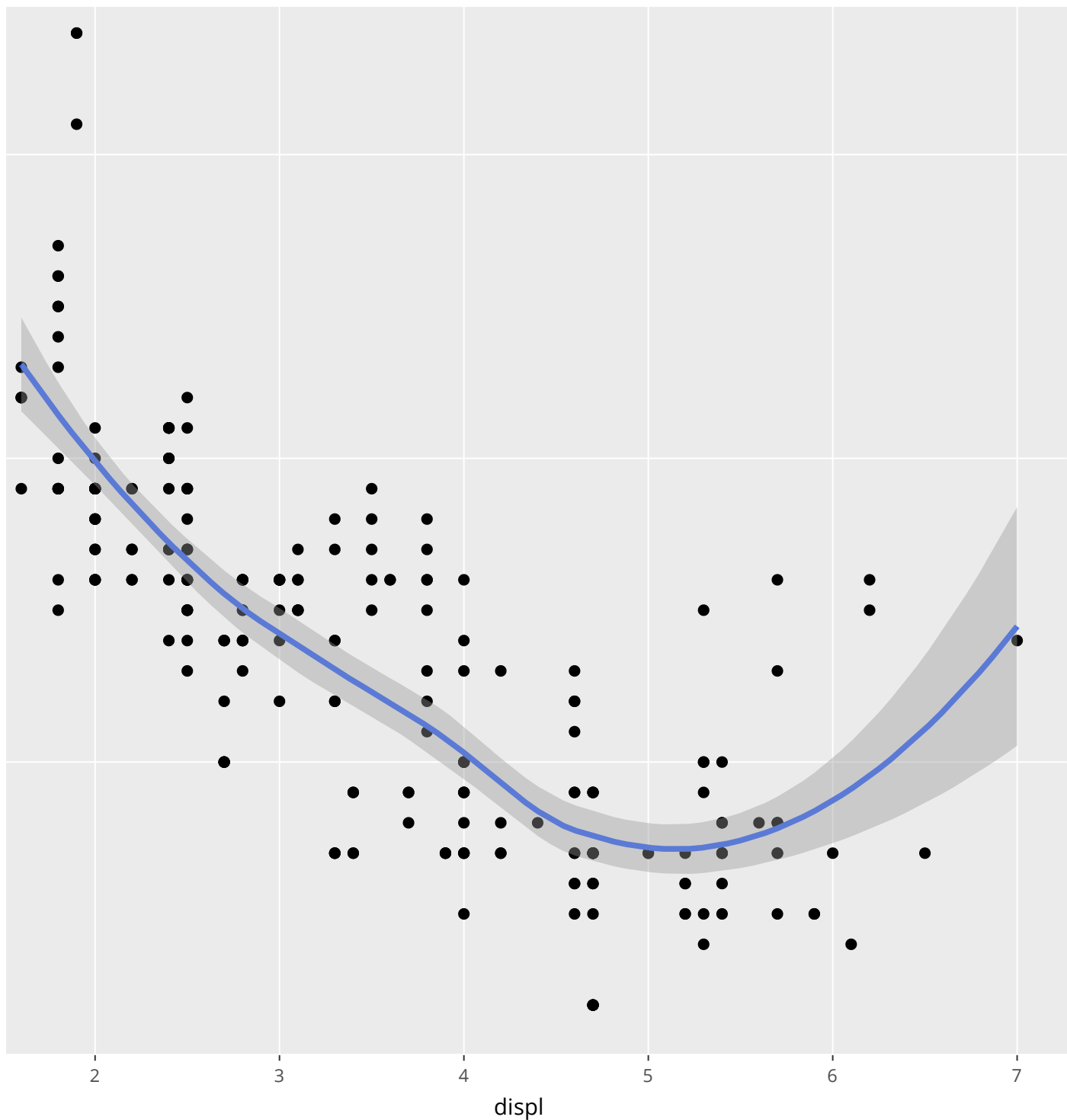
```
om_smooth()` using method = 'loess' and formula 'y ~ x'
```



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

plots may also be produced in the HTML output of the document:

```
library("plotly")  
plotly(p)
```



on Code

Package `reticulate` (<https://cran.r-project.org/package=reticulate>) (Ushey, Allaire, and Tang 2020) includes an engine for R Markdown that enables easy interoperability between Python and R chunks. Below we show a small subset of the available functionalities. We refer to the vignette `R Markdown Python Engine` (https://cran.r-project.org/web/packages/reticulate/vignettes/r_markdown.html) for a more detailed description.

```
/"reticulate")
tualenv("computo-template")
```

be sure (here, in R) that the required python modules are available

```
if(!_module_available("seaborn")) py_install("seaborn")
if(!_module_available("pandas")) py_install("pandas")
if(!_module_available("matplotlib")) py_install("matplotlib")
```

python

copy of python code and associated output:

```
pandas as pd
seaborn as sns
matplotlib.pyplot as plt

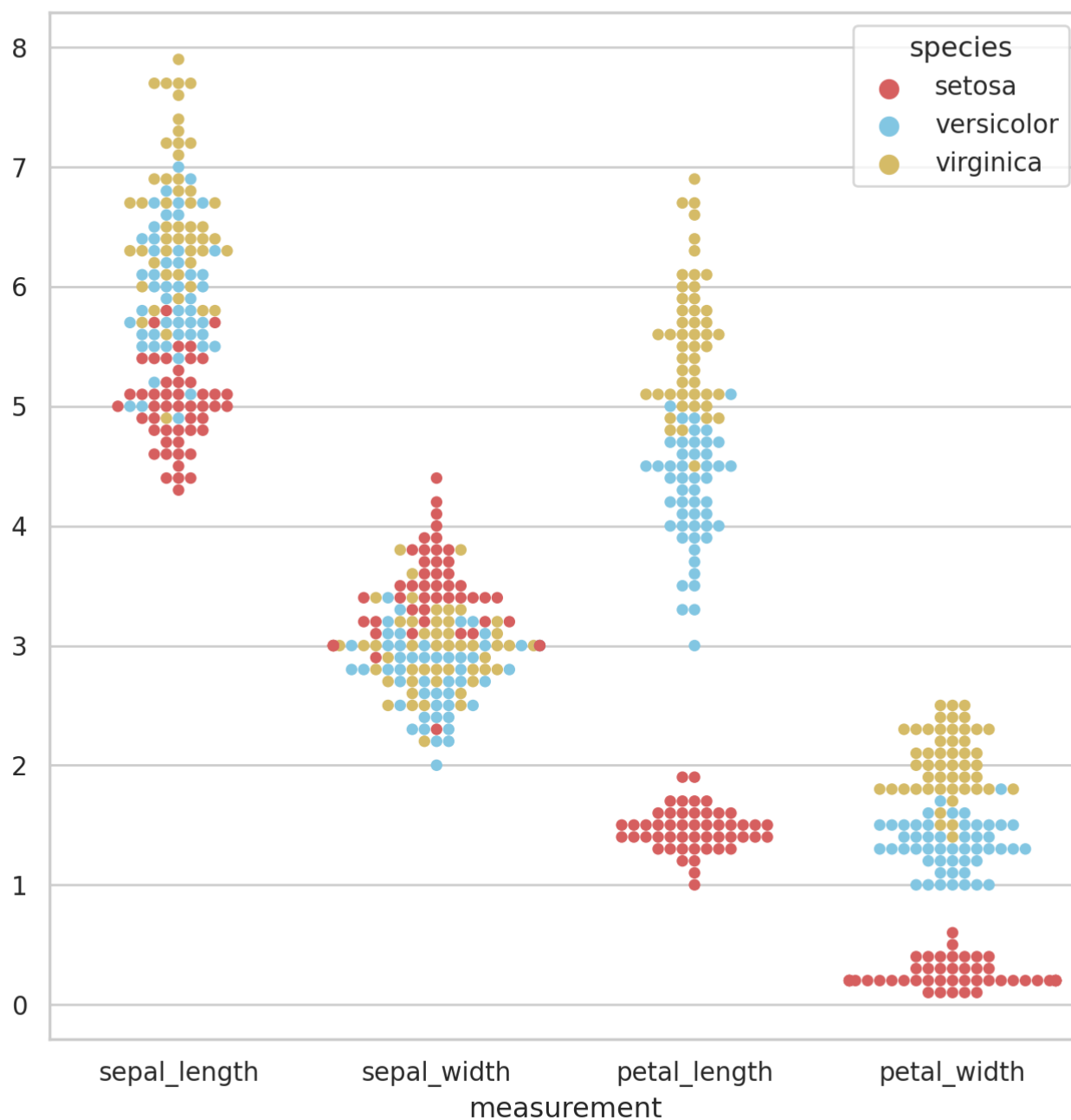
plt.rcParams["style"]="whitegrid", palette="muted")

# load the example iris dataset
sns.load_dataset("iris")

# convert the dataset to "long-form" or "tidy" representation
pd.melt(iris, "species", var_name="measurement")

# create a categorical scatterplot to show each observation
sns.scatterplot(x="measurement", y="value", hue="species", palette=["r", "c", "y"], data=iris)

plt.show()
```



munication between R and python chunks

s created within Python chunks are available to R using the `py` object exported by the reticulate e.g.:

```
own::paged_table(py$iris)
```

	measurement<chr>	value<dbl>
	sepal_length	5.1
	sepal_length	4.9
	sepal_length	4.7
	sepal_length	4.6
	sepal_length	5.0
	sepal_length	5.4
	sepal_length	4.6
	sepal_length	5.0
	sepal_length	4.4
	sepal_length	4.9
500 rows	Previous	123456...60Next

ly, all objects created within R are available from Python using the `r` object exported by the `reticulate` package:

is create an object within R:

```
volcano)

rown::paged_table(as.data.frame(volcano))
```

V1<dbl>	V2<dbl>	V3<dbl>	V4<dbl>	V5<dbl>	V6<dbl>	V7<dbl>	V8<dbl>	V9<dbl>	V10<dbl>
100	100	101	101	101	101	101	100	100	100
101	101	102	102	102	102	102	101	101	101
102	102	103	103	103	103	103	102	102	102
103	103	104	104	104	104	104	103	103	103
104	104	105	105	105	105	105	104	104	103
105	105	105	106	106	106	106	105	105	104
105	106	106	107	107	107	107	106	106	105
106	107	107	108	108	108	108	107	107	106
107	108	108	109	109	109	109	108	108	107
108	109	109	110	110	110	110	109	109	108
37 rows 1-10 of 61 columns	Previous	1	2	3	4	5	6	...	9Next

it is accessible from Python:


```
~.volcano)
```

```
90. 100. 101. ... 104. 104. 103.]
91. 101. 102. ... 105. 104. 104.]
92. 102. 103. ... 105. 105. 104.]
.
98. 98. 98. ... 94. 94. 94.]
97. 98. 98. ... 94. 94. 94.]
97. 97. 97. ... 94. 94. 94.]]
```

r languages

ally, you can include many others languages into Rmarkdown including Julia and C++. If you are
ple enough to configure binder and prove us the reproducibility of your code, feel free to use any
guage.

figuring binder

: binder setup

em libraries

ckages

ient.yml for packages available in conda, or install.R for other packages (including git(hub)
available via remotes or Bioconductor packages availabl via BiocManager).

ling

ommit which include de string “do_build” in its message. (try to lower the footprint...)

ision information

```
Info()
```

```

ersion 4.0.3 (2020-10-10)
iform: x86_64-conda-linux-gnu (64-bit)
ing under: Ubuntu 20.04.2 LTS

ix products: default
3/LAPACK: /usr/share/miniconda/envs/computorbuild/lib/libopenblas-r0.3.12.so

ale:
| LC_CTYPE=en_US.UTF-8      LC_NUMERIC=C
| LC_TIME=en_US.UTF-8      LC_COLLATE=en_US.UTF-8
| LC_MONETARY=en_US.UTF-8  LC_MESSAGES=en_US.UTF-8
| LC_PAPER=en_US.UTF-8     LC_NAME=C
| LC_ADDRESS=C             LC_TELEPHONE=C
| LC_MEASUREMENT=en_US.UTF-8 LC_IDENTIFICATION=C

ached base packages:
 stats      graphics  grDevices  utils      datasets  methods    base

or attached packages:
 reticulate_1.18 plotly_4.9.3    ggplot2_3.3.3

led via a namespace (and not attached):
| tidyselect_1.1.0  xfun_0.20      purrr_0.3.4    splines_4.0.3
| lattice_0.20-41  colorspace_2.0-0 vctrs_0.3.6    generics_0.1.0
| htmltools_0.5.1.1 viridisLite_0.3.0 yaml_2.2.1     mgcv_1.8-34
| utf8_1.1.4       rlang_0.4.10   pillar_1.5.1   glue_1.4.2
| withr_2.4.1      DBI_1.1.1      rappdirs_0.3.3 lifecycle_1.0.0
| stringr_1.4.0    munsell_0.5.0  gtable_0.3.0   htmlwidgets_1.5.3
| evaluate_0.14    labeling_0.4.2 knitr_1.31      crosstalk_1.1.1
| fansi_0.4.2      highr_0.8      Rcpp_1.0.6     scales_1.1.1
| jsonlite_1.7.2   farver_2.1.0   digest_0.6.27  stringi_1.5.3
| dplyr_1.0.5      grid_4.0.3     tools_4.0.3    magrittr_2.0.1
| lazyeval_0.2.2   tibble_3.1.0   crayon_1.4.1   tidyr_1.1.3
| pkgconfig_2.0.3  ellipsis_0.3.1 Matrix_1.3-2    data.table_1.14.0
| assertthat_0.2.1 rmarkdown_2.7  http_1.4.2     R6_2.5.0
| nlme_3.1-152     compiler_4.0.3

```

rences

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

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

John J. Allaire, JJ Allaire, and Yuan Tang. 2020. *Reticulate: Interface to Python*.

<https://github.com/rstudio/reticulate> (<https://github.com/rstudio/reticulate>).