

Shaped around
the individual

Quantitative research

Reproducibility and replication: a guide

Paul Webb, Head of Research, Praxis Care



APA JARS

- APA (2020) p. 77.
- Journal article reporting standards (JARS) e.g.
- apastyle.apa.org/jars
 - Guidelines->Special Designs->Replication Studies
 - JARS-Quant | Table 6

Replication

- Research results may be considered *replicable* if:

“there is sufficient information available for independent researchers to make the same findings using the same procedures with new data.” (Gandrud 2020 p. 4)

- Data *and* code should be accessible to others.
- Research *is* the data *and* the software environment *and* code *and* the method of knowledge transfer.

Why do reproducible research?

Reproducible research:

- Makes assessment of existing research easier
- Is time efficient
- Encourages teamwork
- Encourages university/non-university working
- Means that changes are easier to make
- Makes it easier to re-start projects in the future
- Encourages the clear specification of problems

The Toolkit

There are six components which make up **a reproducible research environment and publisher:**

- R
- Knitr and rmarkdown
- Markdown and/or LaTeX
- Rstudio
- Cloud storage plus version control
- Shell programs

Gandrud 2020 pp 11 – 12.

The Toolkit: Scope

This presentation will focus on:

- R
- Knitr and rmarkdown
- Rstudio

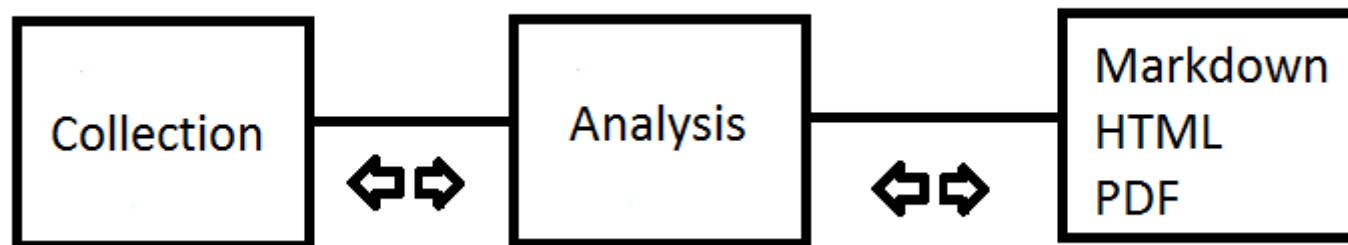
Gandrud 2020 pp 11 – 12.

Why is the toolkit different?

The toolkit is distinctive because:

- Everything is a text file
- Text files are future proof
- Files are human-readable
- It is cross-platform
- It encourages modularity
- It is free/open source

Fitting the pieces together

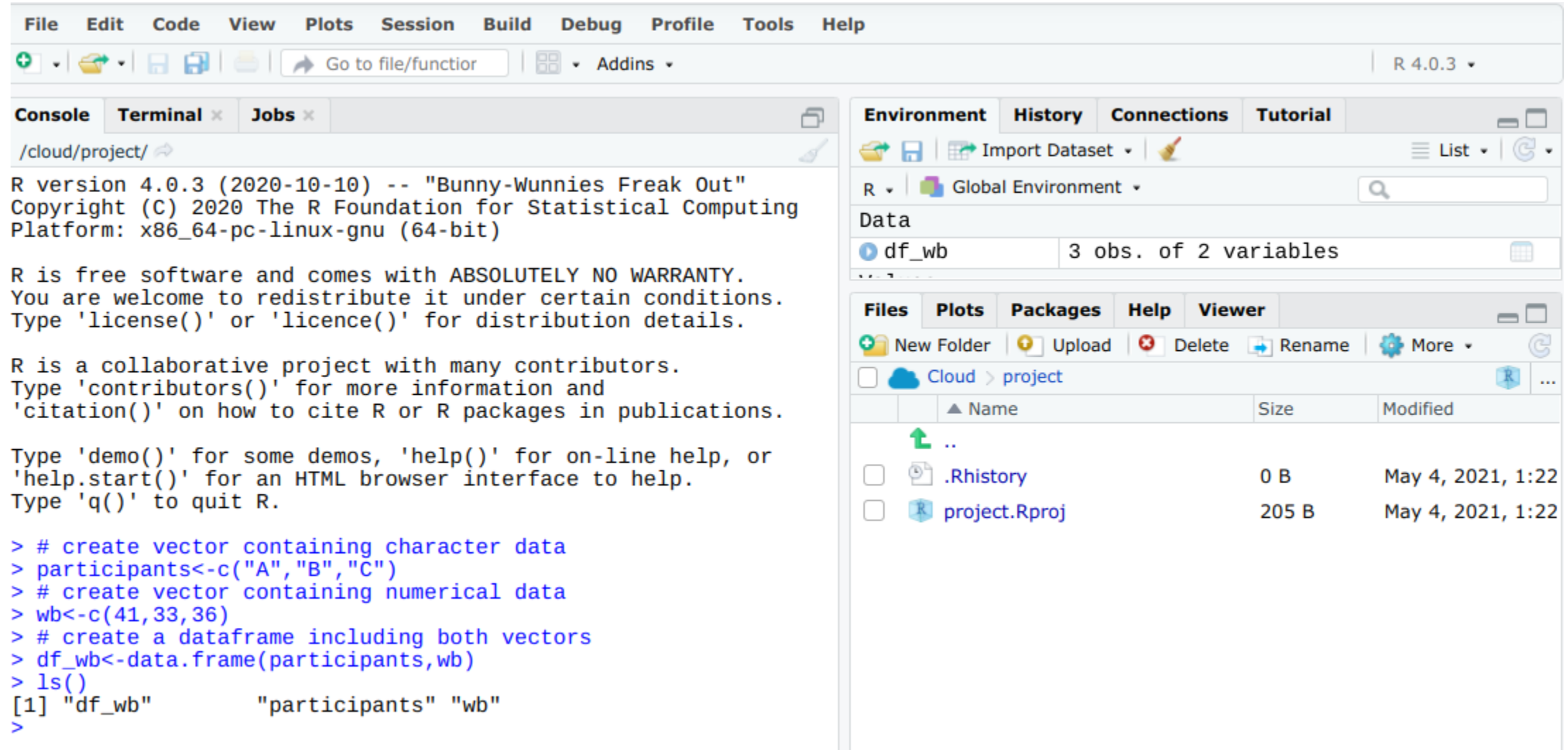


read.table
read.csv
download.file
source
make

R
R packages

knitr
rmarkdown
![alt_text](file_path)
kable
xtable
ggplot2

R via RStudio



The screenshot displays the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, and Help. Below the menu is a toolbar with icons for file operations and a search bar. The main window is divided into several panes:

- Console:** Shows the R version (4.0.3) and the start of a script. The script contains R code to create vectors and a dataframe, followed by the `ls()` command. The output shows the objects `df_wb`, `participants`, and `wb`.
- Environment:** Shows the Global Environment with a search bar and a data frame `df_wb` containing 3 observations of 2 variables.
- Files:** Shows a file browser for the `project` directory. It lists files `..`, `.Rhistory` (0 B), and `project.Rproj` (205 B), all modified on May 4, 2021, at 1:22.

```
R version 4.0.3 (2020-10-10) -- "Bunny-Wunnies Freak Out"
Copyright (C) 2020 The R Foundation for Statistical Computing
Platform: x86_64-pc-linux-gnu (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

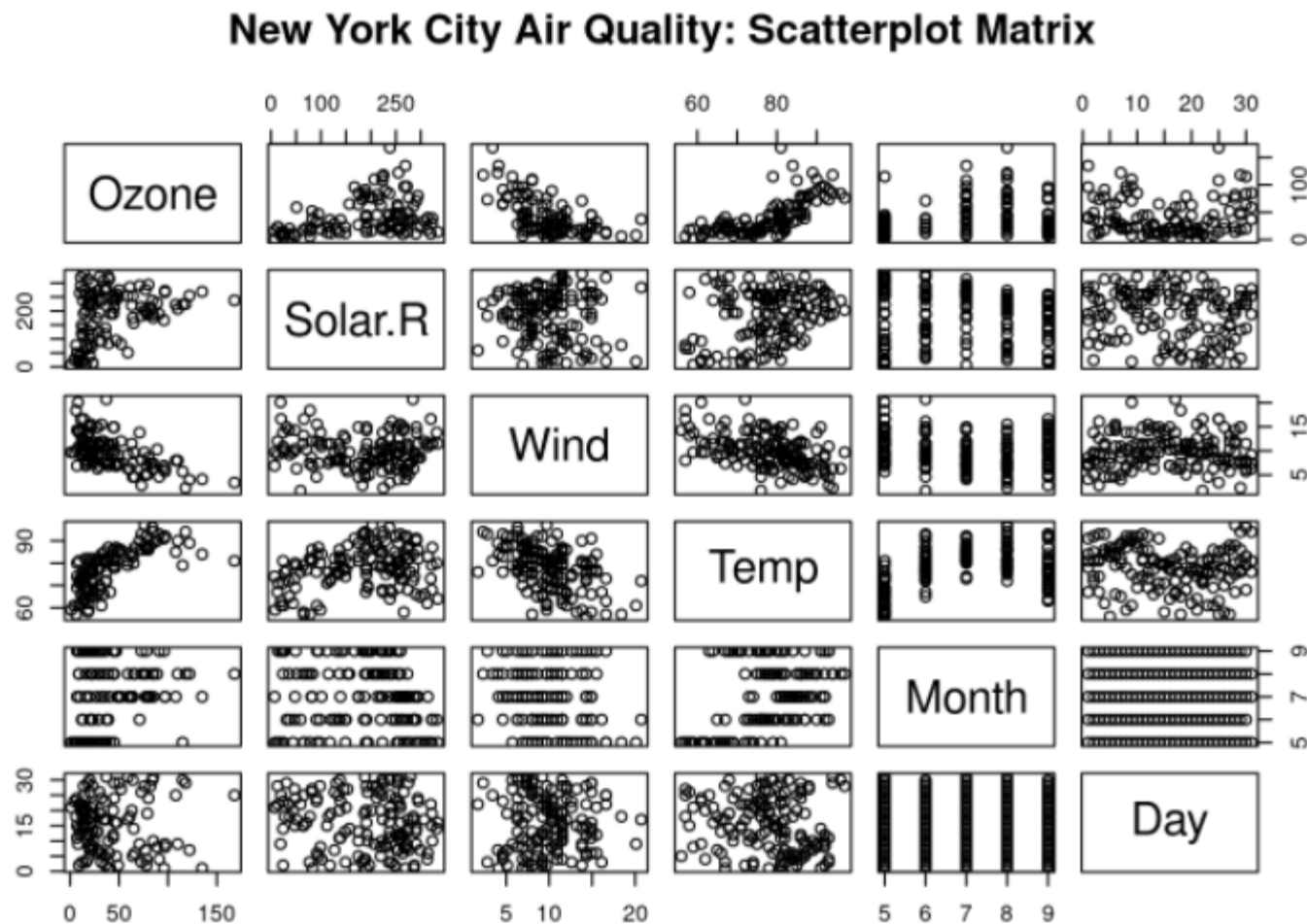
> # create vector containing character data
> participants<-c("A","B","C")
> # create vector containing numerical data
> wb<-c(41,33,36)
> # create a dataframe including both vectors
> df_wb<-data.frame(participants,wb)
> ls()
[1] "df_wb"      "participants" "wb"
>
```

Name	Size	Modified
..		
.Rhistory	0 B	May 4, 2021, 1:22
project.Rproj	205 B	May 4, 2021, 1:22

```
plot(airquality, main = "New York City Air Quality: Scatterplot Matrix")
```

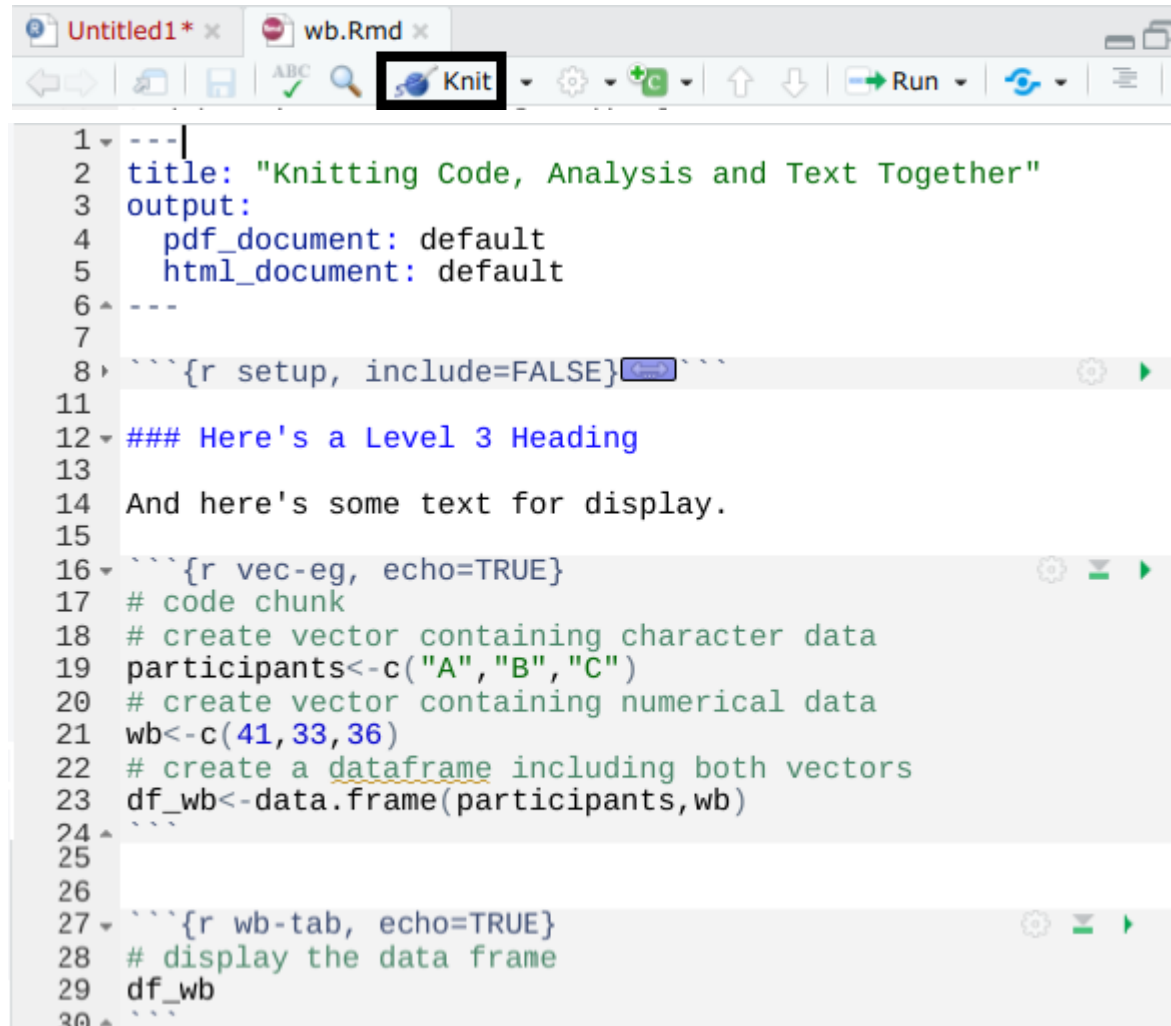
Plotting

```
34 You can also embed plots, for example:
35
36   ```{r airquality, echo=TRUE}
37   plot(airquality, main = "New York City Air Quality:
38   Scatterplot Matrix")
39   ```
```



Rmarkdown

Rmarkdown file



```
1 ---|
2 title: "Knitting Code, Analysis and Text Together"
3 output:
4   pdf_document: default
5   html_document: default
6 ---
7
8 {r setup, include=FALSE}
9
11
12 ### Here's a Level 3 Heading
13
14 And here's some text for display.
15
16 {r vec-eg, echo=TRUE}
17 # code chunk
18 # create vector containing character data
19 participants<-c("A","B","C")
20 # create vector containing numerical data
21 wb<-c(41,33,36)
22 # create a dataframe including both vectors
23 df_wb<-data.frame(participants,wb)
24
25
26
27 {r wb-tab, echo=TRUE}
28 # display the data frame
29 df_wb
30
```

Code chunk visible in PDF

Knitting Code, Analysis and Text Together

Here's a Level 3 Heading

And here's some text for display.

```
# code chunk
# create vector containing character data
participants<-c("A","B","C")
# create vector containing numerical data
wb<-c(41,33,36)
# create a dataframe including both vectors
df_wb<-data.frame(participants,wb)

# display the data frame
df_wb
```

```
## participants wb
## 1           A 41
## 2           B 33
## 3           C 36
```

Code chunk not visible in PDF

Here's a Level 3 Heading

And here's some text for display.

```
## participants wb
## 1           A 41
## 2           B 33
## 3           C 36
```

Resources

- **R:** The R Project for Statistical Computing <https://www.r-project.org/>
- **Knitr:** Report generation with R <https://yihui.org/knitr/>
- **R Markdown:** An Authoring Framework <https://rmarkdown.rstudio.com/>
- **RStudio:** An Integrated Development Environment for R <https://www.rstudio.com/>

Bibliography

APA (2020) *Publication Manual of the American Psychological Association: The Official Guide to APA Style*, 7th edition. Washington DC: APA.

Gandrud, C. (2020) *Reproducible Research with R and RStudio*, 3rd edition. Boca Raton London NY: Chapman and Hall/CRC.

Kabakoff, R.I. (2019) *R in Action: Data Analysis and Graphics with R*. Shelter Island NY: Manning Publications Co.

R Core Team (2020). *R: A language and environment for statistical computing*. Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>

RStudio Team (2021). *RStudio: Integrated Development Environment for R*. RStudio, PBC, Boston, MA URL <http://www.rstudio.com/>.

Xie, Y. (2015) *Dynamic Documents with R and knitr*. 2nd edition. Chapman and Hall/CRC.

Thank you

For further information
contact Paul
paulwebb@praxiscare.org.uk