Shaped around
the individual

Quantitative research

Reproducibility and replication: a guide

Paul Webb, Head of Research, Praxis Care

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:
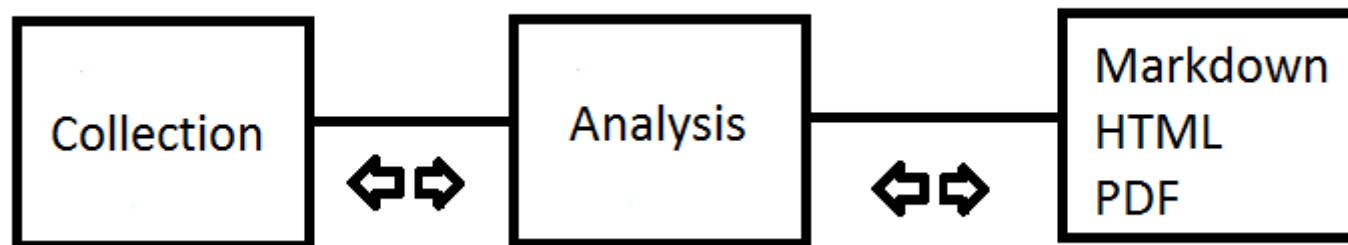
- R

- 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

Praxis Care

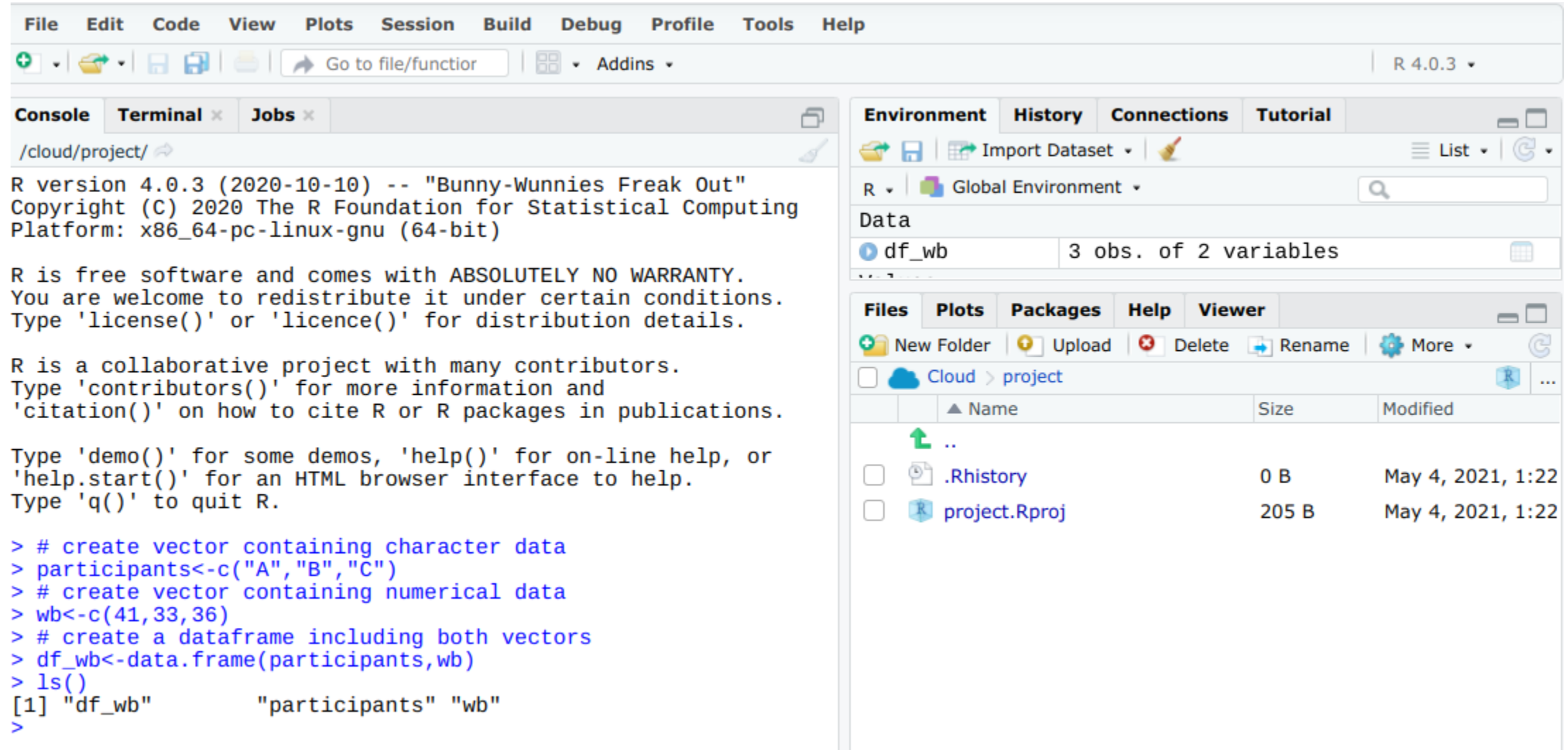# Fitting the pieces together

Praxis Care

```
+-------------+       +-------------+       +-------------+
| Collection  | <==>  |  Analysis   | <==>  | Markdown    |
|             |       |             |       | HTML        |
|             |       |             |       | PDF         |
+-------------+       +-------------+       +-------------+
```

read.table   R   knitr
read.csv   R packages   rmarkdown
download.file
source   kable
make   xtable
                                             ggplot2

R via RStudio

Plotting

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

You can also embed plots, for example:

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

**New York City Air Quality: Scatterplot Matrix**

# Rmarkdown

Praxis Care

## Rmarkdown file

```
Untitled1*     wb.Rmd

ABC   Q    Knit  ▾   ⚙ ▾ ⊕C ▾   ⇧ ⇩   → Run ▾   ⟳ ▾   ≡
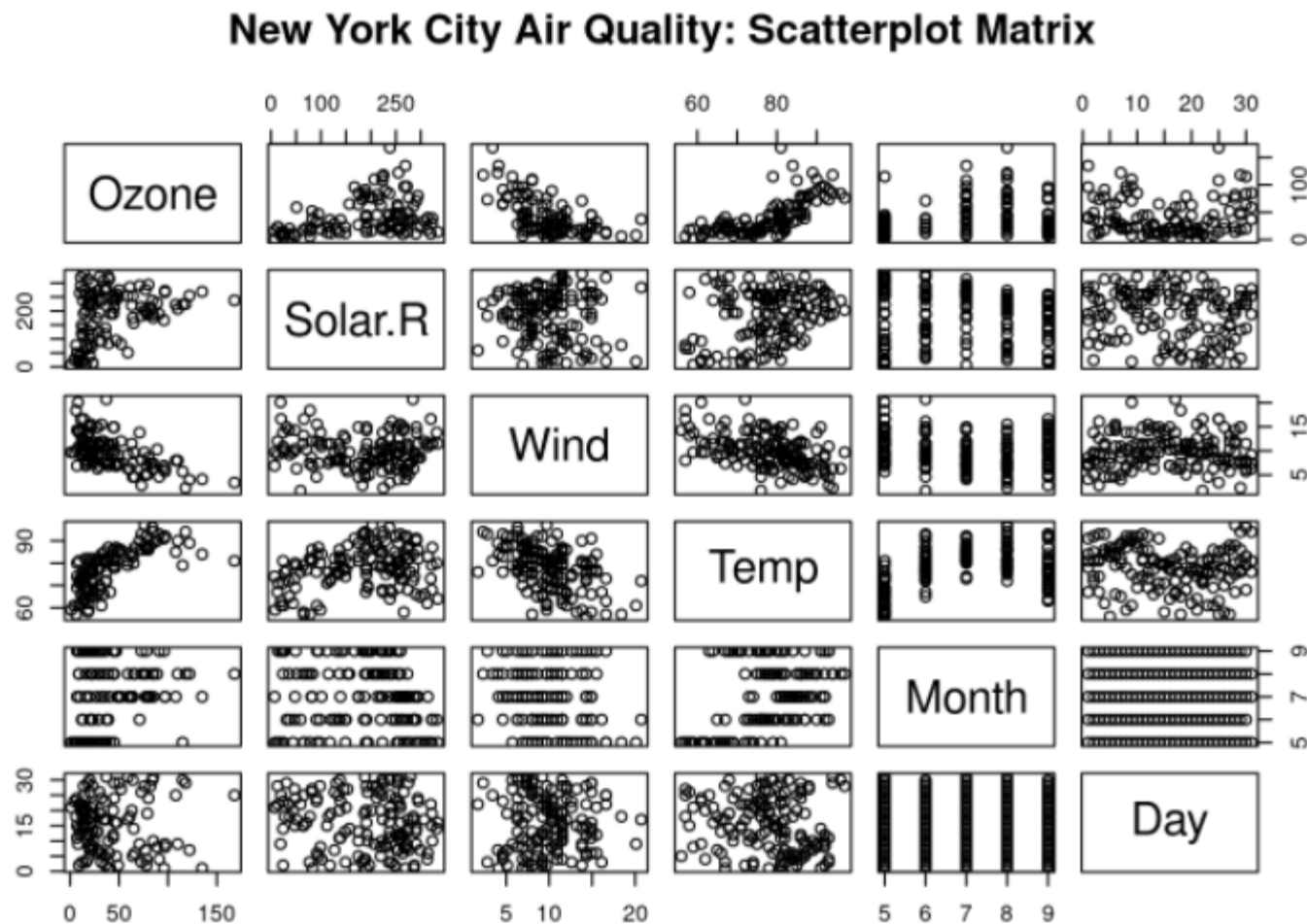
 1  ---
 2  title: "Knitting Code, Analysis and Text Together"
 3  output:
 4    pdf_document: default
 5    html_document: default
 6  ---
 7
 8  ```{r setup, include=FALSE}```
11
12  ### Here's a Level 3 Heading
13
14  And here's some text for display.
15
16  ```{r vec-eg, echo=TRUE}
17  # code chunk
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.

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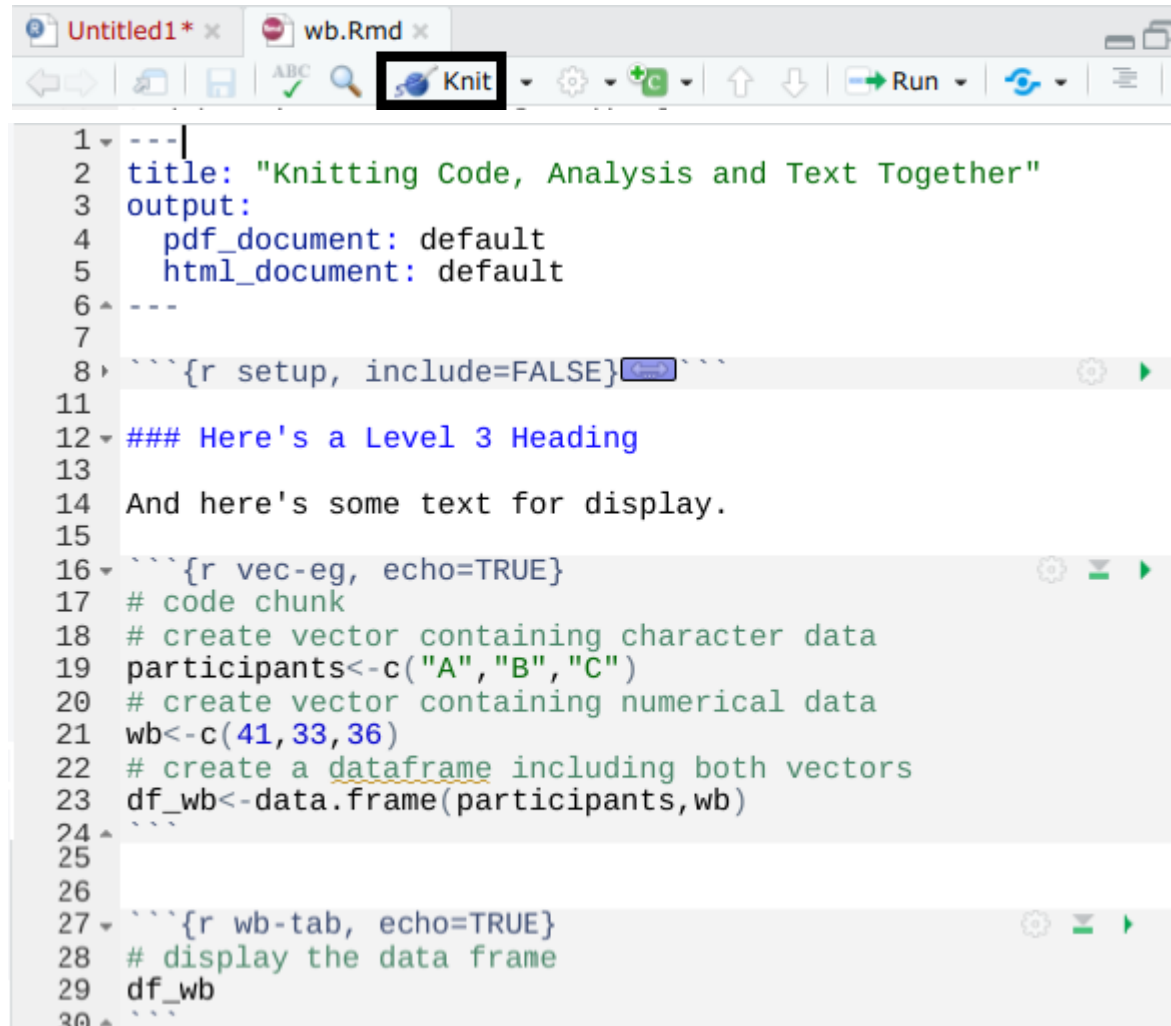# code chunk
# create vector containing character data
participants<-c("A","B","C")
# create vector containing numerical data
wb<-c(41,33,36)
# create a dataframe including both vectors
df_wb<-data.frame(participants,wb)

# 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

Praxis Care

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