

Integrating analysis code and document preparation : An example Rmarkdown + papaja document

Tom Stafford¹ & Mate Giannadou²

¹ Department of Psychology, University of Sheffield

² Department of Psychosocial Science, University of Bergen

This is an example of a note

R is a statistical programming language which is increasingly popular with psychologists. It can import and process your data, fit statistical models (from simple t-tests to state of the art such as bayesian multilevel model fitting). It also makes nice plots. RStudio is a way of editing R scripts and running R analysis. RMarkdown is a way of using RStudio to produce documents (e.g. as webpages, MS Word or PDF). Another advantage is that you can include R code in your document file - so no more running your analysis in SPSS and copying the results into your document (and making errors / forgetting which version of the analysis you ran etc). This is an example document which integrates all the functions of Rmarkdown - running analysis, formatting references, etc. It uses an add-on for Rmarkdown called papaja which helps us make nicely APA formatted documents

Introduction

You will find it useful to compare the output PDF document with the .rmd document. This latter item is the thing you edit to produce the PDF.

Example Subheading

Here are some example references in the following sentence. For reviews of this topic see Wickelgren (1977); Heitz (2014). Here is another example reference. For example, ignoring speed of response is common in psychophysics, whereas some domains of cognitive psychology where high-accuracy is assumed, focus only on response times (e.g. Stafford, Ingram, & Gurney, 2011)¹, albeit sometimes after a cursory check that standard null-hypothesis tests do not reveal significant differences in error-rates.

Method

Rmarkdown also lets us track figure labels, and updates them automatically. Look! Kittens! Illustrated in Figure 1.

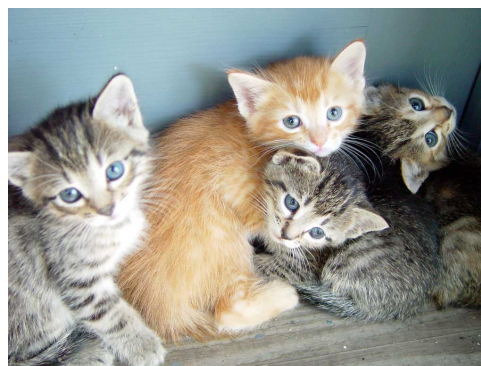


Figure 1. Example figure caption

Requirements

You should install R, RStudio and tex and papaja. More details here https://crsh.github.io/papaja_man/introduction.html#getting-started

Results

Now let's integrate some R code to generate/import some data, run and analyse and integrate it into the document: You can't see it, but inbetween this paragraph and the last we asked R to generate some random data and save it to a CSV file. Now we're going to import the data from the CSV file, as if it was independently created data - from an experiment or similar - and plot a graph.

Correspondence concerning this article should be addressed to Tom Stafford, Department of Psychology, University of Sheffield, 1 Vicar Lane, Sheffield, S1 2LT, UK. E-mail: t.stafford@sheffield.ac.uk

¹As well as an example of a footnote.

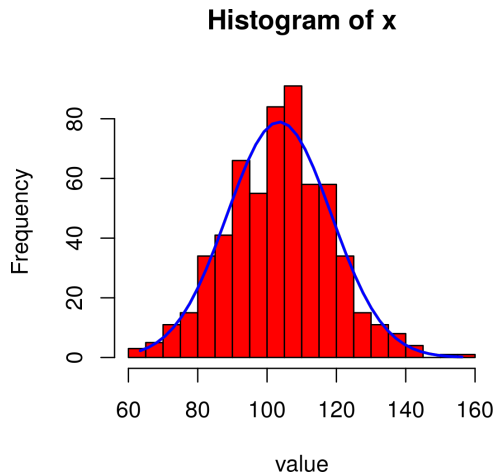


Figure 2. Histogram of all data, grouped

See Figure 2. Of course we could draw all sorts of things, but this is a proof-of-concept. Finally, let's run a t-test and integrate the results into the text.

We found there was a statistically significant difference between the two groups ($t=-6.32$ (591.61), $p = 0.00$). Note how the exact values in the previous sentence change every time we re-make the document (because the document also re-generates the underlying data).

Unanswered questions: Is this the best way to integrate values into text? Why is the df not an integer? What is the best way to define figure sizes so you get nice and/or consistent sizing across document formats?

Discussion

Rmarkdown is good. Need to change reference style? Change one line. Need to submit as PDF rather than .DOC? Just click "Word" as output rather than "PDF" (instructions here https://rmarkdown.rstudio.com/articles_docx.html). Need to change to two column style to make a nice pre-print? Again, simple - just change one line!

Main conclusions

Of course, there's more effort in installing and learning and correctly marking up your document in the first place, but it is worth it

Acknowledgements

This is an example acknowledgements section

References

- Heitz, R. P. (2014). The speed-accuracy tradeoff: History, physiology, methodology, and behavior. *Frontiers in Neuroscience*, 8, 150.
- Stafford, T., Ingram, L., & Gurney, K. N. (2011). Piéron's law holds during stroop conflict: Insights into the architecture of decision making. *Cognitive Science*, 35(8), 1553–1566.
- Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, 41(1), 67–85.