

# Introduction to R Markdown

Mike Smith

April 6, 2018

# What is Markdown?

- Lightweight markup language
- Write text and 'tag' elements for formatting later
  - Similar to LaTeX or HTML, but much simpler
- Can be used to create many types of output

# What is Markdown?

- Lightweight markup language
- Write text and 'tag' elements for formatting later
  - Similar to LaTeX or HTML, but much simpler
- Can be used to create many types of output
- Time for a live demo
  - `example_1-layout.md`

# What is R Markdown?

- Basically **R** + Markdown
- You can include blocks of code

# What is R Markdown?

- Basically **R** + Markdown
- You can include blocks of code
- Allows you to keep in a single document:
  - Code
  - Explanatory text
  - Output e.g. tables, figures

# Vignette examples

# Literate Programming & Live Documents

*“The main idea is to regard a program as a communication to human beings rather than as a set of instructions to a computer.” - Donald Knuth*

- Combines ‘human readable’ prose with source code in a single document
- Analyses can be easily understood and repeated
- Facilitates repeatable research by allowing workflows to be regularly tested for robustness over time

# Repeatability vs Reproducibility

- Can you get my results from my data?



# How can R Markdown help?

- How often do you have code/ideas/figures spread across multiple files?
- An example
  - <https://doi.org/10.1186/1471-2105-11-208>
- Look at Supplementary File 2

# R Code Chunks

- `example_2-codechunks.Rmd`
- `example_3-chunkOptions.Rmd`
- `example_4-simplePipeline.Rmd`

## sessionInfo()

- Just because you can run a script later, doesn't mean it gives the same answers!
  - R changes
  - Packages change
  - Locale settings etc

# Publishing with R Markdown

# Bioconductor Workflows

- Educational documents detailing how to use multiple packages
- Regularly checked on Bioconductor servers
  - Ensures changes in component tools are identified quickly
  - Allows rapid updates by author
  - Provides users a platform to easily access the complete suite of tools
- Difficult to get credit

# F1000Research

- F1000Research provides a citable, peer-reviewed publication platform
- Intention is for the same document to be submitted to both platforms - updates are possible

F1000Research  
PDF

Source  
R Markdown



```
1 # Source: RNA-seq analysis to easy as 1-2-3 with limma, Glimma and edgeR
2 #
3 # Author: Michael Smith
4 #
5 # Affiliation: The author and Eliza Hall Institute of Medical Research, 145 Resnet Parade,
6 # Parkville, VIC 3045, Melbourne, Australia; Department of Medical Biotechnology, The University of
7 # Melbourne, Parkville, VIC 3010, Melbourne, Australia
8 #
9 # Email: Michael.Smith@unimelb.edu.au
10 #
11 # Date: 14 September 2016
12 #
13 # Keywords: RNA-seq analysis, limma, Glimma, edgeR
14 #
15 # License: CC-BY-NC-ND 4.0 International
16 #
17 #
18 #
19 #
20 #
21 #
22 #
23 #
24 #
25 #
26 #
27 #
28 #
29 #
30 #
31 #
32 #
33 #
34 #
35 #
36 #
37 #
38 #
39 #
40 #
41 #
42 #
43 #
44 #
45 #
46 #
47 #
48 #
49 #
50 #
51 #
52 #
53 #
54 #
55 #
56 #
57 #
58 #
59 #
60 #
61 #
62 #
63 #
64 #
65 #
66 #
67 #
68 #
69 #
70 #
71 #
72 #
73 #
74 #
75 #
76 #
77 #
78 #
79 #
80 #
81 #
82 #
83 #
84 #
85 #
86 #
87 #
88 #
89 #
90 #
91 #
92 #
93 #
94 #
95 #
96 #
97 #
98 #
99 #
100 #
```



# BiocWorkflowTools

```
source("https://bioconductor.org/biocLite.R")  
biocLite("BiocWorkflowTools")
```

# Conclusions

- R Markdown is great for:
  - Lab book / journal recording
  - Reports
  - Supplementary materials
- As a principle document for journal articles - we're not there yet!



# Acknowledgements

---

Andrzej Oles  
Wolfgang Huber  
Mike Love  
Bernd Klaus

---

