My Example Computed Manuscript
Created in Rmarkdown

Jeffrey M. Perkel ·

Abstract A mock computed manuscript created in RStudio using {Rmarkdown}. The {Bookdown} and {Rticles} packages were used to output the text in Springer Nature's desired manuscript format.

Keywords

1 Introduction

This is an R Markdown document. Markdown is a simple formatting syntax for authoring HTML, PDF, and MS Word documents. For more details on using R Markdown see `http://rmarkdown.rstudio.com`.

Here we'll add some references from Zotero (Perkel 2020): (Fisch et al. 2015; Argelaguet et al. 2021; Lê Cao et al. 2021).

Markdown documents can include inline equations written in LaTeX, such as $F = ma$. Here is an equation on its own line:

$$a^2 + b^2 = c^2$$

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Jeffrey M. Perkel
Springer Nature, 1 New York Plaza, New York, NY
E-mail: `jeffrey.perkel@nature.com`

Table 1 Original subject data

ID	class	wt	—	ID	class	wt	—	ID	class	wt
1	GRP_3	246	—	34	GRP_3	201	—	67	GRP_3	286
2	GRP_3	128	—	35	GRP_2	241	—	68	GRP_2	246
3	GRP_3	229	—	36	GRP_2	236	—	69	GRP_3	163
4	GRP_3	127	—	37	GRP_1	240	—	70	GRP_3	102
5	GRP_3	163	—	38	GRP_3	116	—	71	GRP_2	276
6	GRP_3	213	—	39	GRP_3	100	—	72	GRP_1	278
7	GRP_3	236	—	40	GRP_3	270	—	73	GRP_1	205
8	GRP_3	136	—	41	GRP_3	286	—	74	GRP_3	76
9	GRP_3	152	—	42	GRP_3	123	—	75	GRP_2	88
10	GRP_3	124	—	43	GRP_3	180	—	76	GRP_1	161
11	GRP_3	215	—	44	GRP_3	176	—	77	GRP_3	116
12	GRP_1	256	—	45	GRP_1	111	—	78	GRP_2	125
13	GRP_2	217	—	46	GRP_3	102	—	79	GRP_3	162
14	GRP_3	218	—	47	GRP_1	289	—	80	GRP_3	238
15	GRP_2	269	—	48	GRP_1	109	—	81	GRP_3	173
16	GRP_3	104	—	49	GRP_3	299	—	82	GRP_3	284
17	GRP_1	97	—	50	GRP_2	273	—	83	GRP_3	296
18	GRP_1	173	—	51	GRP_2	255	—	84	GRP_3	206
19	GRP_3	78	—	52	GRP_1	219	—	85	GRP_3	232
20	GRP_3	274	—	53	GRP_3	167	—	86	GRP_3	218
21	GRP_3	262	—	54	GRP_2	241	—	87	GRP_3	145
22	GRP_3	143	—	55	GRP_1	274	—	88	GRP_3	128
23	GRP_3	112	—	56	GRP_3	269	—	89	GRP_2	131
24	GRP_1	102	—	57	GRP_3	229	—	90	GRP_2	288
25	GRP_3	178	—	58	GRP_1	172	—	91	GRP_1	249
26	GRP_1	269	—	59	GRP_3	190	—	92	GRP_3	173
27	GRP_3	242	—	60	GRP_3	245	—	93	GRP_1	82
28	GRP_2	240	—	61	GRP_1	166	—	94	GRP_1	172
29	GRP_3	147	—	62	GRP_3	141	—	95	GRP_3	102
30	GRP_3	106	—	63	GRP_1	152	—	96	GRP_2	286
31	GRP_1	262	—	64	GRP_3	254	—	97	GRP_1	215
32	GRP_2	289	—	65	GRP_3	170	—	98	GRP_1	116
33	GRP_1	114	—	66	GRP_2	219	—	99	GRP_1	226

2 Results

2.1 Inline computation

One 'killer feature' of computed manuscripts is the ability to compute and insert values into the text rather than requiring authors to input them manually. That circumvents the possibility that the author will enter an incorrect number, or forget to update them should new data arise.

For instance, imagine we are analyzing data from a clinical trial:

We have **99** subjects in our study (Table 1). The average weight is **192.73** (range: **76**-**299**). We have **25** subjects in Group 1, **17** subjects in Group 2, and **57** in Group 3. (The numbers in **bold face type** are computed values.)

Table 2 New subject data

ID	class	wt	—	ID	class	wt	—	ID	class	wt
1	GRP_3	235	—	21	GRP_3	129	—	41	GRP_2	185
2	GRP_2	230	—	22	GRP_3	174	—	42	GRP_2	299
3	GRP_3	281	—	23	GRP_1	97	—	43	GRP_3	214
4	GRP_3	185	—	24	GRP_3	276	—	44	GRP_2	207
5	GRP_1	82	—	25	GRP_3	136	—	45	GRP_1	80
6	GRP_2	173	—	26	GRP_3	291	—	46	GRP_3	95
7	GRP_3	251	—	27	GRP_1	277	—	47	GRP_3	151
8	GRP_1	123	—	28	GRP_2	230	—	48	GRP_3	171
9	GRP_3	242	—	29	GRP_3	277	—	49	GRP_3	257
10	GRP_3	195	—	30	GRP_2	241	—	50	GRP_1	237
11	GRP_1	233	—	31	GRP_3	251	—	51	GRP_1	103
12	GRP_1	298	—	32	GRP_3	291	—	52	GRP_3	214
13	GRP_1	217	—	33	GRP_1	99	—	53	GRP_3	234
14	GRP_1	236	—	34	GRP_1	260	—	54	GRP_1	255
15	GRP_3	284	—	35	GRP_1	88	—	55	GRP_1	257
16	GRP_1	263	—	36	GRP_1	161	—	56	GRP_2	116
17	GRP_1	250	—	37	GRP_1	75	—	57	GRP_3	284
18	GRP_3	123	—	38	GRP_1	247	—	58	GRP_3	165
19	GRP_3	249	—	39	GRP_3	191	—	59	GRP_1	299
20	GRP_3	286	—	40	GRP_3	184	—	60	GRP_2	144

Now suppose we get another tranche of data:

There are **60** subjects in this new dataset (Table 2). Their average weight is **206.3** (range: **75**-**299**).

Combining the two datasets, we have a total of **159** subjects. The revised average weight is **197.85** (range: **75**-**299**). We now have **47** subjects in Group 1, **26** subjects in Group 2, and **86** in Group 3.

2.2 The Fibonacci sequence

As Rmarkdown documents can do anything R can do, we can also create and include figures. For instance, we can calculate the first **25** numbers in the Fibonacci sequence. The data are shown in Figure 1. Note that this figure number (as well as the table numbers above) are automatically generated.

3 Methods

The following code was used in section 2.1 to create the original clinical trial data:

```
# create a mock dataset
n_subjs <- 99
subjID <- 1:n_subjs
```
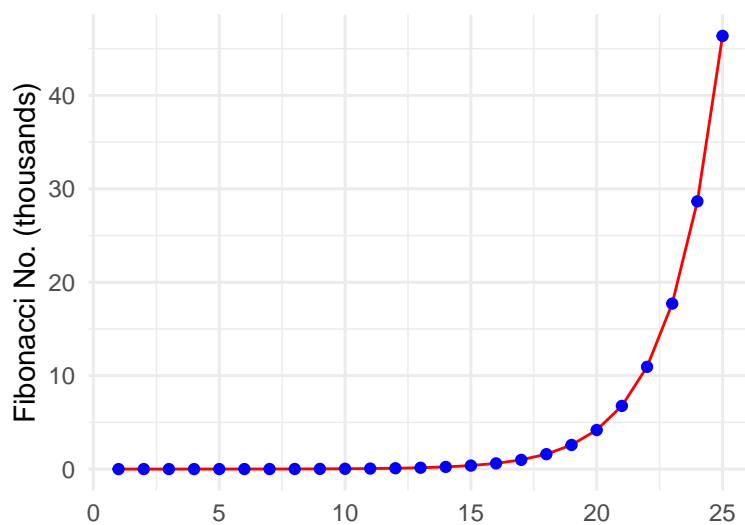
Fig. 1 The first 26 numbers of the Fibonacci sequence

```r
# generate 99 random #s between 1 and 10
tmp <- floor(runif(n_subjs, min = 1, max = 10))
# assign those numbers to any of 3 subject groups
fn <- function(x) {
  if (x > 7) 'GRP_1'
  else if (x > 5) 'GRP_2'
  else 'GRP_3'
}
subj_class <- sapply(tmp, fn)

# pick random weights between 75 and 300
wts <- floor(runif(n_subjs, min = 75, max = 300))
# combine the data into a table
df <- data.frame(ID = subjID, class = subj_class, wt = wts)

# display the table, splitting the 99 rows into 3 cols wide
tmp <- cbind(df[1:33,], rep('|', 33),
             df[34:66,], rep('|', 33),
             df[67:99,])
names(tmp) <- c('ID', 'class', 'wt', '|', 'ID', 'class', 'wt',
                '|', 'ID', 'class', 'wt')
knitr::kable(tmp, format = 'latex', booktabs = TRUE,
             caption = "Original subject data")
```

The following code was used in section 2.2 to calculate and plot the Fibonacci sequence:

```r
# calculate the first `n` Fibonacci numbers
fibonacci <- function(n) {
  stopifnot(n > 2)
  ar <- rep(0, n)
  ar[1] <- f1 <- 0
  ar[2] <- f2 <- 1
  for (i in 3:n) {
    f3 <- f1 + f2
    ar[i] <- f3
    f1 <- f2
    f2 <- f3
  }
  return (ar)
}

ncalc <- 25
f <- fibonacci(ncalc)
df <- data.frame(1:ncalc, f)
names(df) <- c("x","y")

div <- 1000

p <- ggplot(df) +
  geom_line(aes(x, y/div), color = "red") +
  geom_point(aes(x, y/div), color = "blue") +
  scale_y_continuous(limits = c(0, max(df$y)/div)) +
  xlab("") +
  ylab("Fibonacci No. (thousands)") +
  theme_minimal()
p
```

References

Argelaguet, Ricard, Anna S. E. Cuomo, Oliver Stegle, and John C. Marioni. 2021. "Computational Principles and Challenges in Single-Cell Data Integration." *Nature Biotechnology*, May, 1–14. https://doi.org/10.1038/s41587-021-00895-7.

Fisch, K. M., T. Meissner, L. Gioia, J.-C. Ducom, T. M. Carland, S. Loguercio, and A. I. Su. 2015. "Omics Pipe: A Community-Based Framework for Reproducible Multi-Omics Data Analysis." *Bioinformatics* 31 (11): 1724–28. https://doi.org/10.1093/bioinformatics/btv061.

Lê Cao, Kim-Anh, Al J. Abadi, Emily F. Davis-Marcisak, Lauren Hsu, Arshi Arora, Alexis Coullomb, Atul Deshpande, et al. 2021. "Community-Wide Hackathons to Identify Central Themes in Single-Cell Multi-Omics." *Genome Biology* 22 (1): 220. `https://doi.org/10.1186/s13059-021-02433-9`.

Perkel, Jeffrey M. 2020. "Streamline Your Writing and Collaborations with These Reference Managers." *Nature* 585 (7823): 149–50. `https://doi.org/10.1038/d41586-020-02491-2`.