

My Example Computed Manuscript

Created in Rmarkdown

Jeffrey M. Perkel ·

Received: date / Accepted: date

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 L^AT_EX, such as $F = ma$. Here is an equation on its own line:

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

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 man-

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	Group 2	252	—	34	Group 3	219	—	67	Group 3	132
2	Group 2	123	—	35	Group 1	218	—	68	Group 3	139
3	Group 2	174	—	36	Group 1	274	—	69	Group 1	286
4	Group 2	206	—	37	Group 3	158	—	70	Group 2	103
5	Group 3	151	—	38	Group 2	216	—	71	Group 3	239
6	Group 1	210	—	39	Group 3	161	—	72	Group 2	129
7	Group 2	264	—	40	Group 3	103	—	73	Group 1	181
8	Group 1	222	—	41	Group 3	194	—	74	Group 1	229
9	Group 3	278	—	42	Group 1	297	—	75	Group 1	260
10	Group 2	218	—	43	Group 2	132	—	76	Group 1	119
11	Group 1	133	—	44	Group 3	88	—	77	Group 3	199
12	Group 2	214	—	45	Group 2	127	—	78	Group 2	113
13	Group 3	126	—	46	Group 2	94	—	79	Group 1	155
14	Group 1	283	—	47	Group 3	82	—	80	Group 1	180
15	Group 2	176	—	48	Group 1	179	—	81	Group 3	181
16	Group 2	195	—	49	Group 1	191	—	82	Group 2	123
17	Group 3	224	—	50	Group 2	147	—	83	Group 3	199
18	Group 2	90	—	51	Group 2	205	—	84	Group 1	215
19	Group 1	197	—	52	Group 2	140	—	85	Group 1	216
20	Group 1	138	—	53	Group 1	253	—	86	Group 3	269
21	Group 1	170	—	54	Group 3	141	—	87	Group 3	249
22	Group 2	232	—	55	Group 2	245	—	88	Group 3	239
23	Group 1	251	—	56	Group 3	116	—	89	Group 2	200
24	Group 3	239	—	57	Group 3	213	—	90	Group 1	191
25	Group 3	248	—	58	Group 1	120	—	91	Group 1	212
26	Group 2	92	—	59	Group 2	197	—	92	Group 3	110
27	Group 1	123	—	60	Group 3	224	—	93	Group 1	86
28	Group 1	90	—	61	Group 2	115	—	94	Group 1	173
29	Group 2	148	—	62	Group 2	197	—	95	Group 3	211
30	Group 3	294	—	63	Group 1	291	—	96	Group 2	133
31	Group 2	219	—	64	Group 3	214	—	97	Group 3	165
32	Group 1	298	—	65	Group 3	281	—	98	Group 3	144
33	Group 1	222	—	66	Group 2	205	—	99	Group 1	280

ually. 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 grouped subjects in three bins and measured their weights.

Rather than analyzing those data and then copying the results into our manuscript, we can use the programming language R to do that in the manuscript itself. Simply enclose the code inside backticks: ``r <code to evaluate>``. For instance, to calculate the circumference and area of a circle with radius $r = 10$, you could write “`A = `r pi * r^2``” and “`C = `r 2 * pi * r``”. Those give $A = 314.159$ and $C = 62.832$.

Table 2 New subject data

ID	class	wt	—	ID	class	wt	—	ID	class	wt
1	Group 3	136	—	21	Group 3	197	—	41	Group 2	247
2	Group 3	227	—	22	Group 2	98	—	42	Group 3	173
3	Group 3	102	—	23	Group 2	296	—	43	Group 1	236
4	Group 2	255	—	24	Group 3	90	—	44	Group 1	240
5	Group 1	191	—	25	Group 1	96	—	45	Group 1	216
6	Group 2	194	—	26	Group 3	124	—	46	Group 3	117
7	Group 3	246	—	27	Group 2	90	—	47	Group 2	202
8	Group 3	280	—	28	Group 2	293	—	48	Group 3	220
9	Group 3	134	—	29	Group 2	259	—	49	Group 1	191
10	Group 2	76	—	30	Group 2	200	—	50	Group 3	129
11	Group 2	154	—	31	Group 2	169	—	51	Group 3	198
12	Group 3	197	—	32	Group 2	138	—	52	Group 2	97
13	Group 1	141	—	33	Group 2	224	—	53	Group 1	281
14	Group 1	110	—	34	Group 2	155	—	54	Group 3	241
15	Group 3	216	—	35	Group 1	81	—	55	Group 2	209
16	Group 2	148	—	36	Group 3	216	—	56	Group 1	151
17	Group 2	188	—	37	Group 2	91	—	57	Group 2	160
18	Group 3	119	—	38	Group 1	270	—	58	Group 2	216
19	Group 3	199	—	39	Group 1	256	—	59	Group 1	269
20	Group 2	124	—	40	Group 3	219	—	60	Group 3	139

We have **99** subjects in our study (Table 1). The average weight is **187.85** (range: **82-298**). We have **34** subjects in Group 1, **32** subjects in Group 2, and **33** in Group 3. (The numbers in **bold face type** are computed values.)

Now suppose we get another tranche of data in the form of a second CSV file:

There are **60** subjects in this new dataset (Table 2). Their average weight is **182.18** (range: **76-296**).

Combining the two datasets, we have a total of **159** subjects. The revised average weight is **185.71** (range: **76-298**). We now have **48** subjects in Group 1, **56** subjects in Group 2, and **55** in Group 3.

2.2 Plotting the data

As Rmarkdown documents can do anything R can do, we can also create and include figures. Here we plot boxplots of our clinical trial data. 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 read in the original clinical trial data from CSV files (these data were simulated):

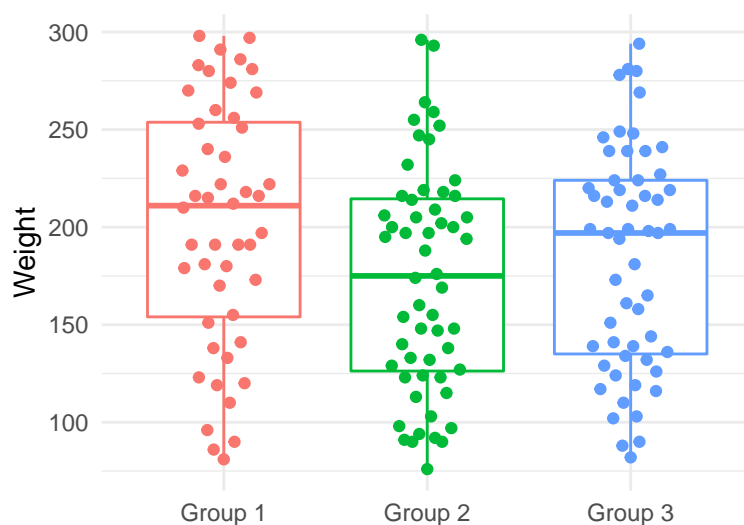


Fig. 1 Weight distribution of clinical trial subjects

```
# read in some example data
df <- read_csv("example-data-1.csv")

# 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 plot the data:

```
final_data %>%
  ggplot(aes(x = class, y = wt, color = class)) +
  geom_boxplot() +
  ggbeeswarm::geom_quasirandom(width = 0.25) +
  xlab("") +
  ylab("Weight") +
  theme_minimal() +
  theme(legend.position = "none")
```

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>.