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.

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_2	216	—	34	GRP_1	80	—	67	GRP_3	193
2	GRP_3	193	—	35	GRP_3	174	—	68	GRP_3	259
3	GRP_1	126	—	36	GRP_1	159	—	69	GRP_3	245
4	GRP_3	113	—	37	GRP_3	183	—	70	GRP_1	254
5	GRP_3	154	—	38	GRP_3	210	—	71	GRP_1	164
6	GRP_2	103	—	39	GRP_3	230	—	72	GRP_3	235
7	GRP_1	278	—	40	GRP_3	162	—	73	GRP_2	266
8	GRP_3	160	—	41	GRP_3	104	—	74	GRP_3	112
9	GRP_1	222	—	42	GRP_3	192	—	75	GRP_1	121
10	GRP_2	242	—	43	GRP_3	281	—	76	GRP_3	172
11	GRP_3	242	—	44	GRP_1	142	—	77	GRP_3	132
12	GRP_1	230	—	45	GRP_3	259	—	78	GRP_3	241
13	GRP_2	254	—	46	GRP_2	125	—	79	GRP_1	294
14	GRP_3	175	—	47	GRP_1	238	—	80	GRP_2	225
15	GRP_2	231	—	48	GRP_3	213	—	81	GRP_3	131
16	GRP_1	132	—	49	GRP_3	163	—	82	GRP_3	99
17	GRP_3	213	—	50	GRP_3	154	—	83	GRP_3	287
18	GRP_3	93	—	51	GRP_3	175	—	84	GRP_2	217
19	GRP_1	212	—	52	GRP_2	188	—	85	GRP_3	245
20	GRP_3	91	—	53	GRP_3	290	—	86	GRP_3	228
21	GRP_3	229	—	54	GRP_3	160	—	87	GRP_3	136
22	GRP_3	284	—	55	GRP_1	115	—	88	GRP_1	211
23	GRP_3	198	—	56	GRP_3	231	—	89	GRP_1	109
24	GRP_1	110	—	57	GRP_1	275	—	90	GRP_2	178
25	GRP_1	226	—	58	GRP_1	255	—	91	GRP_3	129
26	GRP_3	202	—	59	GRP_3	164	—	92	GRP_3	94
27	GRP_3	163	—	60	GRP_3	195	—	93	GRP_3	262
28	GRP_2	106	—	61	GRP_2	250	—	94	GRP_2	281
29	GRP_1	207	—	62	GRP_3	80	—	95	GRP_3	279
30	GRP_2	140	—	63	GRP_3	280	—	96	GRP_3	216
31	GRP_3	151	—	64	GRP_1	117	—	97	GRP_3	267
32	GRP_3	110	—	65	GRP_2	280	—	98	GRP_1	136
33	GRP_3	214	—	66	GRP_1	97	—	99	GRP_2	93

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 **189.41** (range: **80**-**294**). 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_2	231	—	21	GRP_2	275	—	41	GRP_3	243
2	GRP_1	140	—	22	GRP_2	181	—	42	GRP_3	251
3	GRP_3	91	—	23	GRP_1	104	—	43	GRP_3	246
4	GRP_3	150	—	24	GRP_2	150	—	44	GRP_3	273
5	GRP_2	276	—	25	GRP_3	157	—	45	GRP_2	226
6	GRP_3	265	—	26	GRP_3	110	—	46	GRP_3	285
7	GRP_3	257	—	27	GRP_2	262	—	47	GRP_3	183
8	GRP_3	284	—	28	GRP_3	201	—	48	GRP_3	285
9	GRP_1	261	—	29	GRP_2	175	—	49	GRP_3	265
10	GRP_1	139	—	30	GRP_3	218	—	50	GRP_3	234
11	GRP_3	262	—	31	GRP_1	129	—	51	GRP_1	96
12	GRP_1	225	—	32	GRP_2	142	—	52	GRP_3	118
13	GRP_1	207	—	33	GRP_3	159	—	53	GRP_2	269
14	GRP_2	80	—	34	GRP_1	216	—	54	GRP_2	205
15	GRP_3	243	—	35	GRP_1	220	—	55	GRP_3	162
16	GRP_3	94	—	36	GRP_2	299	—	56	GRP_2	247
17	GRP_3	270	—	37	GRP_3	278	—	57	GRP_3	296
18	GRP_3	235	—	38	GRP_3	284	—	58	GRP_2	195
19	GRP_3	283	—	39	GRP_3	122	—	59	GRP_3	180
20	GRP_3	209	—	40	GRP_2	126	—	60	GRP_1	128

Now suppose we get another tranche of data:

There are **60** subjects in this new dataset (Table 2). Their average weight is **206.62** (range: **80-299**).

Combining the two datasets, we have a total of **159** subjects. The revised average weight is **195.91** (range: **80-299**). We now have **36** subjects in Group 1, **33** subjects in Group 2, and **90** 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 create the original clinical trial data:

```r
# create a mock dataset
n_subjs <- 99
subjID <- 1:n_subjs
```
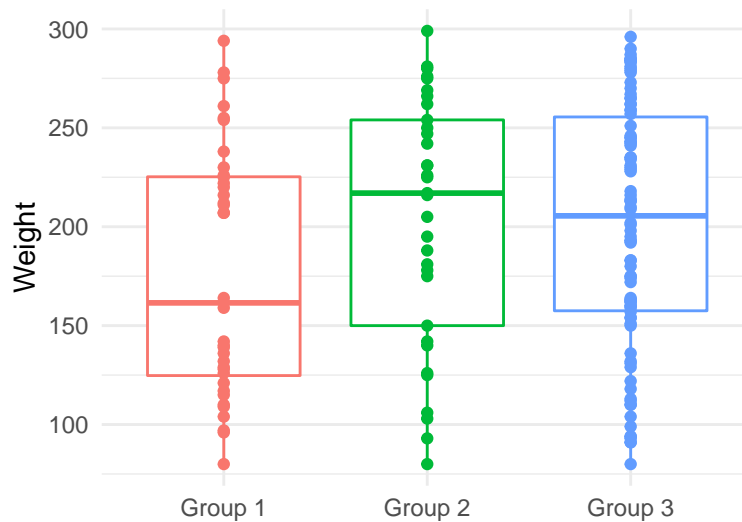
Fig. 1 Weight distribution of clinical trial subjects

```
# 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 plot the data:

```
final_data %>%
  mutate(class = str_replace(class, '_', ' '),
         class = str_replace(class, 'GRP', 'Group')) %>%
  ggplot(aes(x = class, y = wt, color = class)) +
  geom_boxplot() +
  geom_point() +
  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.