





Public Benefit Corporation

Create free and open-source software for data science, scientific research, and technical communication

J.J. Allaire CEO, RStudio rstudio.com/pbc-keynote

Details of Certified B Corporation assessment at <https://bcorporation.net/directory/rstudio>



abbvie

Johnson & Johnson

 NOVARTIS

 Pfizer

 gsk

 Roche

Genentech

We're proud of our partnerships, helping drive scientific innovation and solve common challenges in pharma



Data Science Journey

Hypothesis
Hardening

Data &

Rich Collaboration

Development

Rapid Iteration

Automation

Data Science Journey

Early Stage
Research

Experiment
Design

Trials

Reporting and
Submission

Hypothesis
Hardening

Data &

Rich Collaboration

Development

Rapid Iteration

Automation

Data Science Journey

Manufacturing
or Marketing
Objective

A/B Test
Predictive Model

Production
Scoring

Hypothesis
Hardening

Data &

Rich Collaboration

Development

Rapid Iteration

Automation

2020 RStudio Pharma Updates

Python &
Kubernetes
support in
RStudio

Python support
in RStudio
Connect
(dash/streamlit/flask)

Python &
BioConductor
in Package
Manager

Hypothesis
Hardening

Data &

Rich Collaboration

Development

Rapid Iteration

Automation

Reimagined R
Markdown

tidymodels

gt package

thematic

Validation
Guidance

Validation Guidance Documents



Shiny and R Markdown

RStudio packages for web applications and reporting



tidyverse, tidymodels, r-lib, gt

RStudio affiliated analysis packages that share validated design principles and practices



RStudio Team

The premiere software bundle for data science teams

- Based on the R core validation guidance, peer reviewed by some of you and RStudio's engineers
- Goal is to position some packages as trusted sources
- And emphasize the benefits of open source software for trustworthy and reproducible science

BioConductor in Package Manager Demo

 Studio Package Manager

R Repos

 all

 cran

Bioconductor Repos

 bioconductor

Info

 Usage Stats

 Admin Guide

 Server API Guide

 Terms and Conditions

bioconductor

Packages

Activity

Setup

What's New Bioconductor and PyPI (Beta)

BIOCONDUCTOR
3.8 (R 3.5)

CLIENT OS
Red Hat Enterprise Linux 7

BioManager will look for Bioconductor packages. If this doesn't show a URL like `https://packagemanager.rstudio.com/bioconductor`, or if you'd like to change it, read on.

Configure the Bioconductor Repository

Bioconductor repository setup code

```
options(BioC_mirror = "https://packagemanager.rstudio.com/bioconductor")
```

We recommend adding this to your R startup file (`Rprofile.site` or `.Rprofile`) to maintain the configuration across R sessions. More information about managing R startup files is available in [this article](#).

Using a CRAN Snapshot

For past Bioconductor releases, we recommend using a CRAN snapshot that provides compatible CRAN packages for that Bioconductor release. Run `BioManager::version()` in an R console to identify your version of Bioconductor, then select an appropriate CRAN snapshot below.

CRAN repository setup code for Bioconductor 3.8 (R 3.5) [change](#)

```
options(repos = c(CRAN = "https://packagemanager.rstudio.com/cran/2018-11-01"))
```

You can also choose to use source or binary CRAN packages. 

Source

Binary

Like the `BioC_mirror` option, you can add this to your R startup file (`Rprofile.site` or `.Rprofile`) to maintain the configuration across R sessions.

gt Demo

```
19 kable <- gt(data, groupname_col = "group") %>%
20   fmt<
21   rows = group == "Completed",
22   columns = c("Placebo", "Low Dose", "High Dose", "Total"),
23   fns = function(x) {sprintf("%d (%g%%)", x, round(x/254*100,0))}
24 ) %>%
25   summary_rows(
26     groups = c("Early Termination"),
27     columns = c("Placebo", "Low Dose", "High Dose", "Total"),
28     fns = list(
29       'Total' = ~sum(., na.rm = TRUE)
30     ),
31     formatter = fmt_number,
32     decimals = 0
33 ) %>%
34   grand_summary_rows(
35     columns = c("Placebo", "Low Dose", "High Dose", "Total"),
36     fns = list(
37       'Total in Trial' = ~sum(., na.rm = TRUE)
38     ),
39   )
```

- Write code once

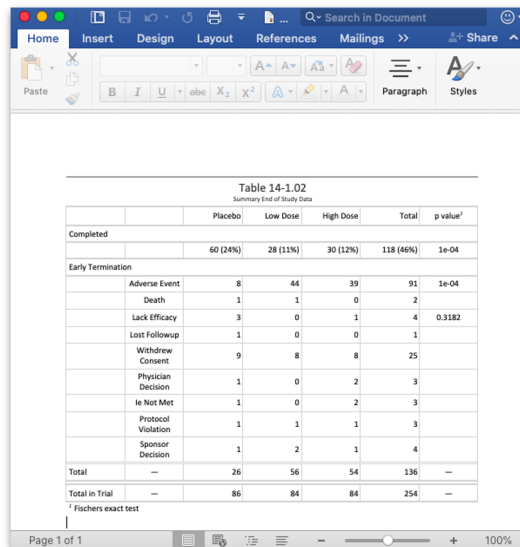


Table 14-1.02
Summary End of Study Data

	Placebo	Low Dose	High Dose	Total	p value [†]
Completed	60 (24%)	28 (11%)	30 (12%)	118 (46%)	1e-04
Early Termination					
Adverse Event	8	44	39	91	1e-04
Death	1	1	0	2	
Lack Efficacy	3	0	1	4	0.3182
Lost Followup	1	0	0	1	
Withdrew Consent	9	8	8	25	
Physician Decision	1	0	2	3	
le Not Met	1	0	2	3	
Protocol Violation	1	1	1	3	
Sponsor Decision	1	2	1	4	
Total	—	26	56	84	—
Total in Trial	—	86	84	170	—

[†] Fischers exact test

- Render to RTF
(for current work)

Table 14-1.02
Summary End of Study Data

	Placebo	Low Dose	High Dose	Total	p value [†]
Completed	60 (24%)	28 (11%)	30 (12%)	118 (46%)	1e-04
Early Termination					
Adverse Event	8	44	39	91	1e-04
Death	1	1	0	2	
Lack Efficacy	3	0	1	4	0.3182
Lost Followup	1	0	0	1	
Withdrew Consent	9	8	8	25	
Physician Decision	1	0	2	3	
le Not Met	1	0	2	3	
Protocol Violation	1	1	1	3	
Sponsor Decision	1	2	1	4	
Total	—	26	56	84	—
Total in Trial	—	86	84	170	—

[†] Fischers exact test

- Render to HTML
(for the shiny future)

WYISWYM Demo

relational-data.Rmd

Knit
Format
Insert
Table

#filtering-joins

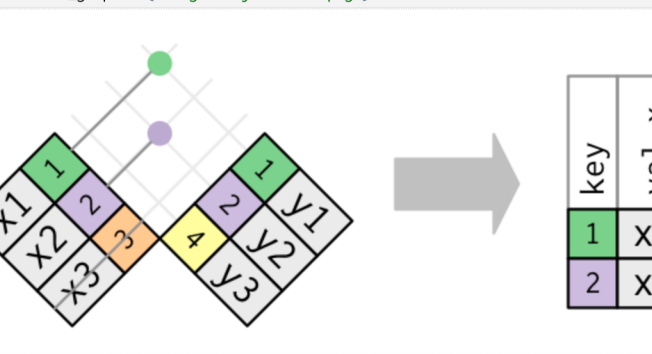
Filtering joins

Filtering joins match observations in the same way as [mutating joins](#), but affect the observations, not the variables¹. There are two types:

<code>semi_join(x, y)</code>	$x \ltimes y$	Keeps all observations in <code>x</code> that have a match in <code>y</code>
<code>anti_join(x, y)</code>	$x \triangleright y$	Drops all observations in <code>x</code> that have a match in <code>y</code>

Graphically, a semi-join looks like this:

```
{r, echo = FALSE, out.width = NULL}
knitr::include_graphics("diagrams/join-semi.png")
```



Only the existence of a match is important; it doesn't matter which observation is matched. This means that filtering joins never duplicate rows like mutating joins do:

References

Validation documents available at the [bottom of this page](#)

Follow the [RStudio blog](#) for the RStudio Package Manager Release with Bioconductor support or visit the [Public RStudio Package Manager](#)

[gt package with RTF support](#)