Bericht zum HaNS-Hackathon

2025-12-05

Table of contents

1 Setup

1.1 Libs

1.2 Other setup

```
source("_common.r")
list.files("funs", full.names = TRUE) |>
  purrr::walk(source)
```

```r
options(digits = 3)
options(tinytable_tt_digits = 2)
```

1.3 Load Targets

```r
tar_load(c(data_prepped, time_spent, course_and_uni_per_visit))
```

2 Challenge 6

```r
n_actions_searches_interactions <-
  data_prepped |>
  select(
    idvisit,
    fingerprint,
    any_of(c(
      "searches",
      "actions",
      "interactions",
      "referrertype",
      "referrername",
      "language",
      "devicetype",
      "devicemodel",
      "operatingsystem",
      "browsername"
    ))
  )
```

2.1 Berechnen Sie die Anzahl der Aktionen pro Visit und pro fingerprint

2.1.a idvisit und fingerprint jeweils unique

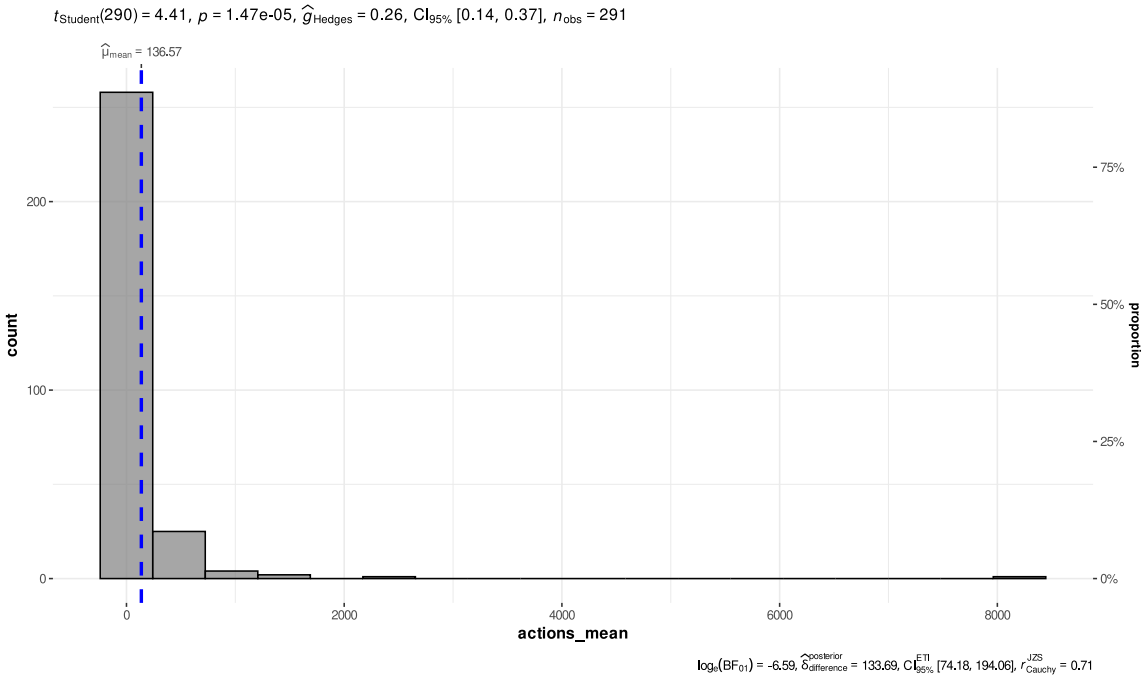
```r
n_actions_searches_interactions |>
  as.data.frame() |>
  summarise(
    idvisit_n = length(unique(idvisit)),
    fingerprint_n = length(unique(fingerprint)),
    actions_mean = mean(as.integer(actions), na.rm = TRUE),
    searches_mean = mean(as.integer(searches), na.rm = TRUE)
  ) |>
  gt()
```

idvisit_n	fingerprint_n	actions_mean	searches_mean
291	127	137	0.206

2.1.b Actions pro idvisit und pro fingerprint

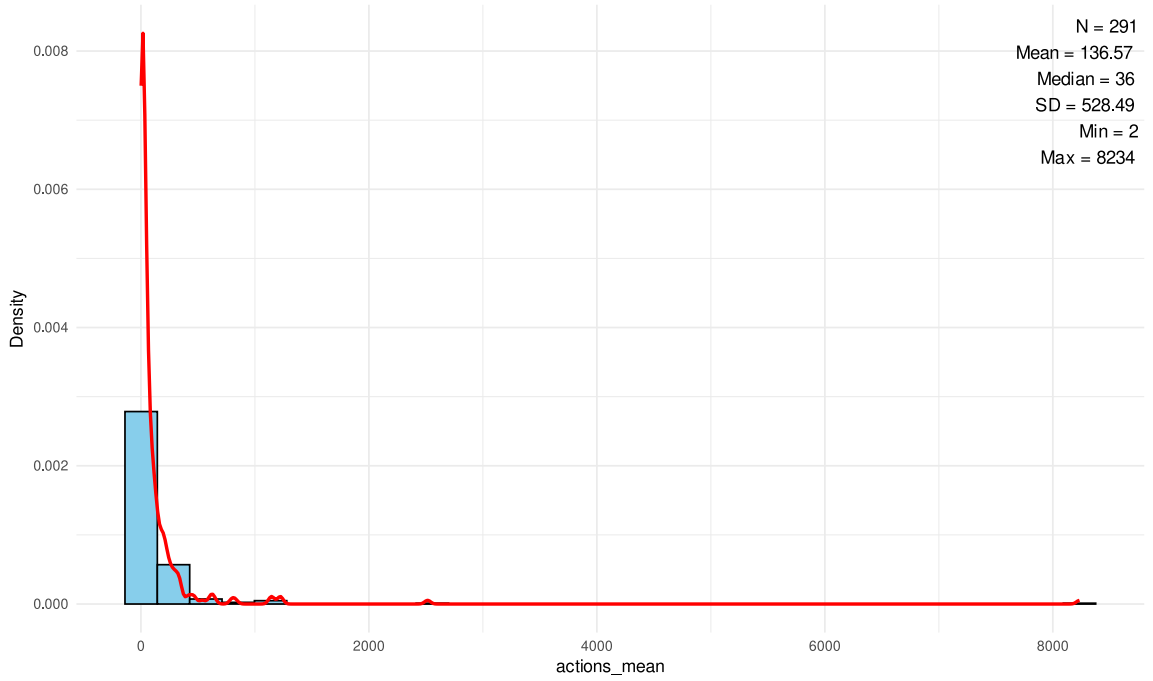
```
n_actions_searches_interactions_summarized <-
n_actions_searches_interactions |>
  group_by(idvisit) |>
  summarise(actions_mean = mean(as.integer(actions), na.rm = TRUE))

n_actions_searches_interactions_summarized |>
  gghistostats(x = actions_mean)
```



https://indrajeetpatil.github.io/ggstatsplot/reference/gghistostats.html

```
n_actions_searches_interactions_summarized |>
  plot_hist_descriptive(actions_mean)
```



> **Note**
>
> Es gibt etwa doppelt so viele Besucher wie unique Nutzer.

2.2 User Specs

2.2.a Referrer Type pro Visit

```
n_actions_searches_interactions |>
  count(referrertype, sort = TRUE)
```

```
   referrertype     n
         <char> <int>
1:      website   251
2:       direct    40
```

2.2.b Referrer Type Name pro Visit

```
n_actions_searches_interactions |>
  count(referrername, sort = TRUE) |>
  knitr::kable()
```

referrername	n
elearning.ohmportal.de	206
NA	40
moodle.hswt.de	30
statics.teams.cdn.office.net	14
hans.th-nuernberg.de	1

2.2.c devicemodel

```
n_actions_searches_interactions |>
  count(devicemodel, sort = TRUE) |>
  slice_head(n = 10)
```

```
                      devicemodel     n
                           <char> <int>
 1:                Generic Desktop   236
 2:                           iPad    25
 3:                  Galaxy A25 5G    11
 4:                         iPhone     8
 5: Galaxy Tab S6 Lite 10.4"" WiFi     4
 6:     Galaxy Tab S8+ 12.4"" WiFi     2
 7:          Generic Mobile device     2
 8:                  Galaxy A15 5G     1
 9:                  Galaxy A32 5G     1
10:                     Galaxy S23     1
```

2.2.d operatingsystem

```
n_actions_searches_interactions |>
  count(operatingsystem, sort = TRUE) |>
  slice_head(n = 10) |>
  tinytable::tt()
```

operatingsystem	n
Windows 11	94
Mac 10.15	76
Windows 10	46
iOS 18.5	27
Mac 15.5	18
Android 15.0	14

operatingsystem	n
Android 13.0	5
GNU/Linux	3
iOS 18.1	3
Android 14.0	1

2.2.e browsername

```
n_actions_searches_interactions |>
  count(browsername, sort = TRUE) |>
  slice_head(n = 10) |>
  tinytable::tt()
```

browsername	n
Safari	74
Chrome	64
Microsoft Edge	63
Mobile Safari	27
Opera GX	21
Chrome Webview	14
Firefox	13
Chrome Mobile iOS	3
Google Search App	3
Opera	3

Die Mac-User scheinen besonders aktiv zu sein auf HaNS.

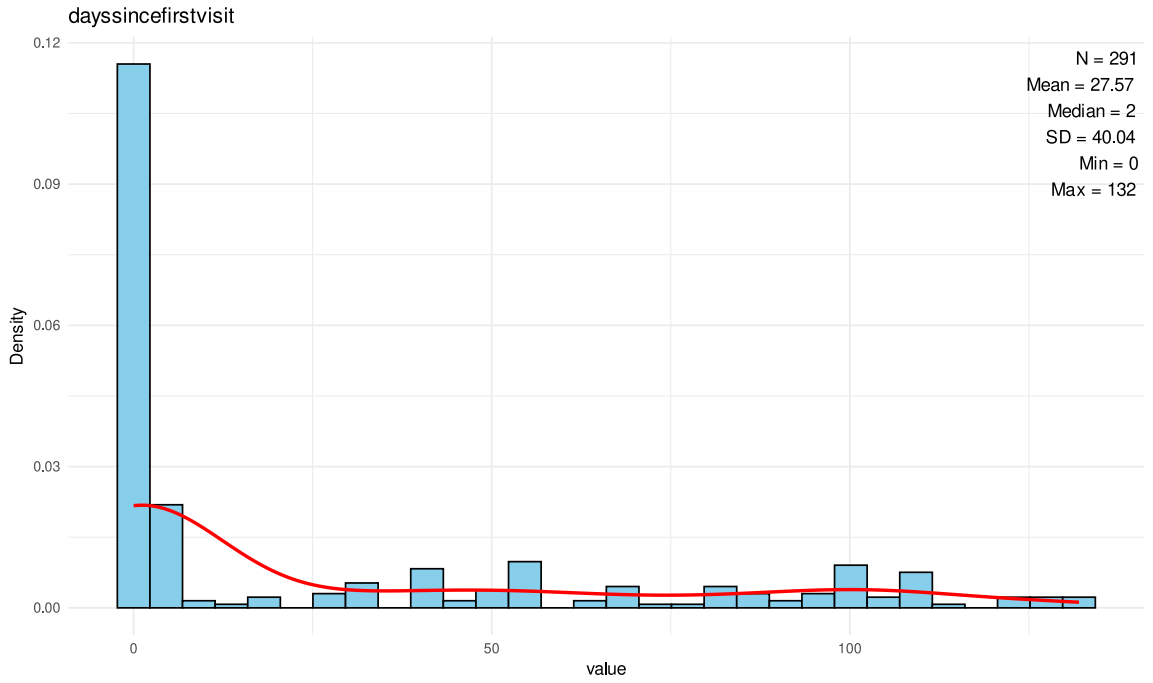
2.3 Berechnen Sie die Anzahl der Tage seit dem letzten Besuch (pro Visit und pro uniquen Besucher)

```
data_prepped |>
  select(matches("days")) |>
  mutate(across(everything(), ~ as.integer(.))) |>
  describe_distribution() |>
  tinytable::tt()
```

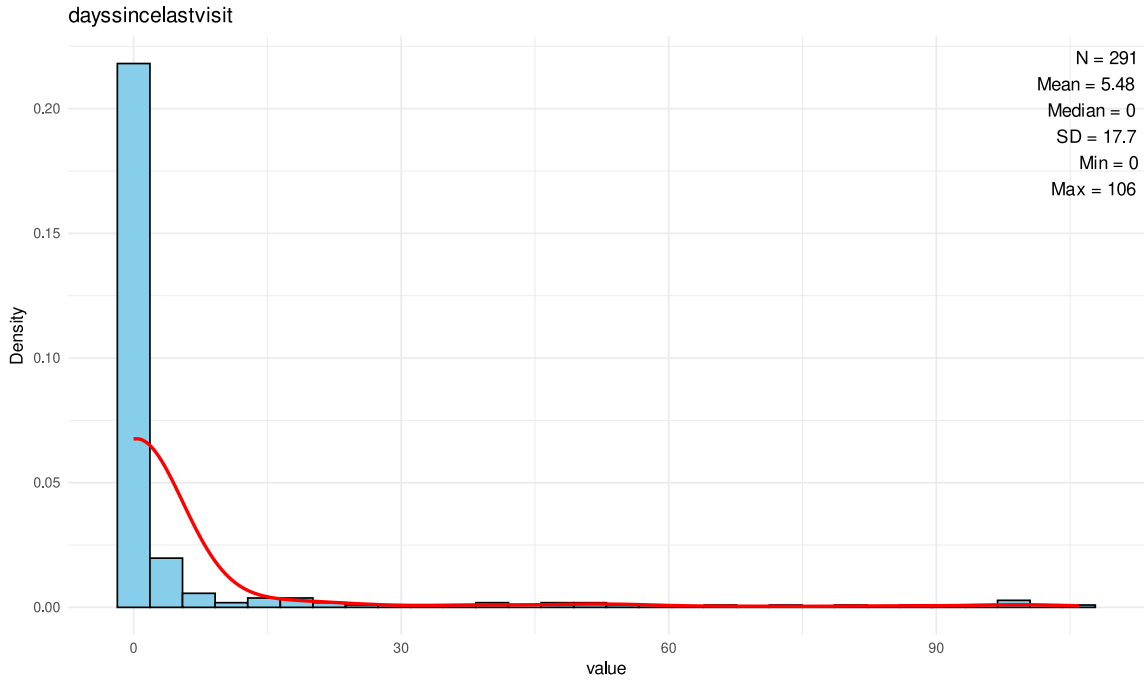
Variable	Mean	SD	IQR	Min	Max	Skewness	Kurtosis	n	n_Missing
dayssince-firstvisit	27.6	40	54	0	132	1.2	−0.062	291	0
dayssince-lastvisit	5.5	18	0	0	106	4	16.294	291	0

```
data_prepped |>
  select(matches("days")) |>
  mutate(across(everything(), as.integer)) |>
  imap(~ plot_hist_descriptive(tibble(value = .x), var = value, title = .y))
```

```
$dayssincefirstvisit
```




```
$dayssincelastvisit
```

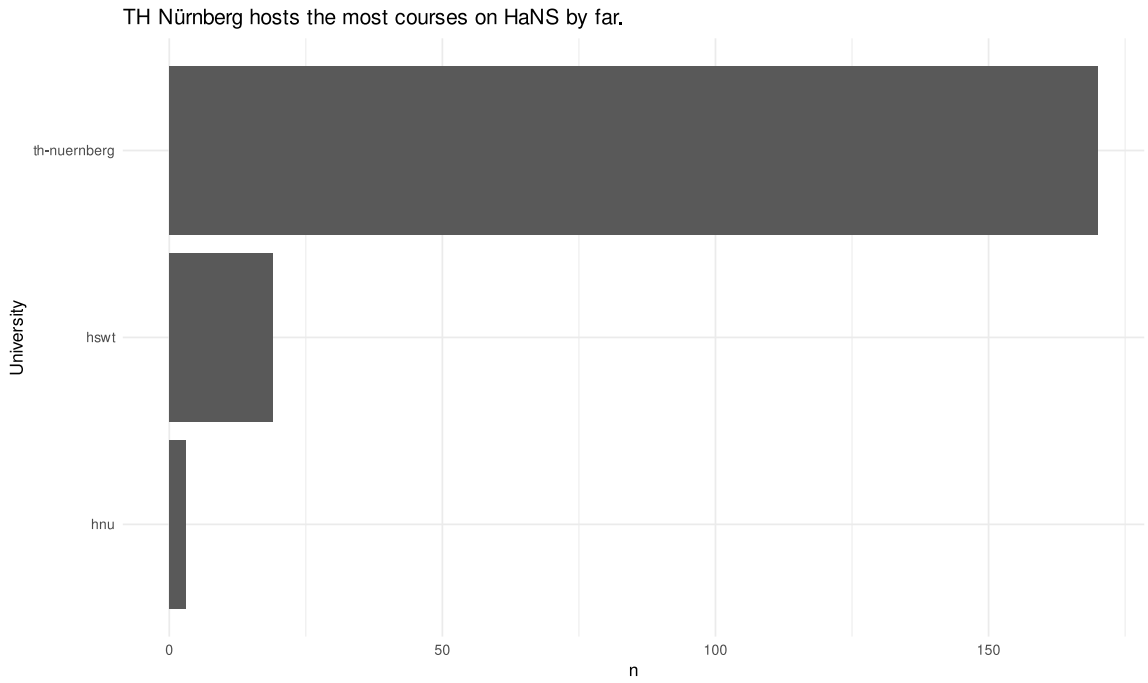


2.4 Anzahl von Unis und Kursen

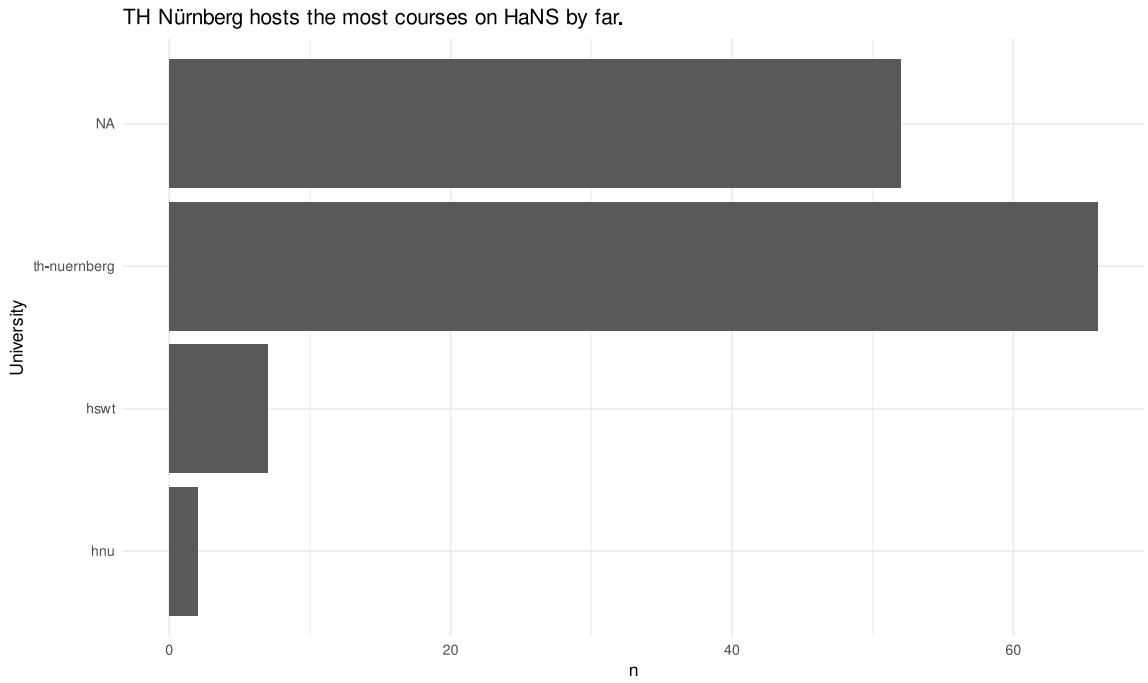
```
course_and_uni_per_visit |>
  count(university) |>
  tinytable::tt()
```

university	n
hnu	3
hswt	19
th-nuernberg	170
NA	99

```
course_and_uni_per_visit |>
  count(university) |>
  drop_na() |>
  ggplot(aes(y = reorder(university, n), x = n)) +
  geom_col() +
  theme_minimal() +
  labs(
    title = "TH Nürnberg hosts the most courses on HaNS by far.",
    y = "University"
  )
```




```
course_and_uni_per_visit |>
  distinct(fingerprint, .keep_all = TRUE) |>
  count(university) |>
  ggplot(aes(y = reorder(university, n), x = n)) +
  geom_col() +
  theme_minimal() +
  labs(
    title = "TH Nürnberg hosts the most courses on HaNS by far.",
    y = "University"
  )
```



2.5 Berechnen Sie die Anzahl der Tage seit dem letzten Besuch pro Modul/ Lehrveranstaltung (pro Visit und pro uniquen Besucher).

Pro idvisit:

```
course_and_uni_per_visit_dayssincelastvisit <-
course_and_uni_per_visit |>
  select(idvisit, course, university) |>
  mutate(idvisit = as.integer(idvisit)) |>
  left_join(data_prepped |> select(idvisit, dayssincelastvisit), by =
"idvisit") |>
  mutate(dayssincelastvisit = as.integer(dayssincelastvisit))
```

```
course_and_uni_per_visit_dayssincelastvisit |>
  group_by(course) |>
  describe_distribution(dayssincelastvisit) |>
  print_md()
```

course	Variable	Mean	SD	IQR	Range	Skew-ness	Kurto-sis	n	n_Missing
bare	dayssince-lastvisit	28.00	27.06	54.00	(0.00, 54.00)	−0.33	−1.50	3	0

course	Variable	Mean	SD	IQR	Range	Skew-ness	Kurto-sis	n	n_Missing
bio	dayssince-lastvisit	1.50	2.12	3.00	(0.00, 3.00)	0.00	−2.00	2	0
cta1	dayssince-lastvisit	6.84	14.19	8.00	(0.00, 58.00)	2.96	9.75	19	0
fodesoa	dayssince-lastvisit	17.60	36.87	19.75	(0.00, 100.00)	1.91	2.30	10	0
fosaq	dayssince-lastvisit	5.80	7.95	14.50	(0.00, 15.00)	0.62	−3.28	5	0
gdi	dayssince-lastvisit	35.33	61.20	106.00	(0.00, 106.00)	1.73	−1.50	3	0
gesoa	dayssince-lastvisit	2.86	11.32	0.00	(0.00, 99.00)	5.88	40.97	142	0
nlp	dayssince-lastvisit	0.00	0.00	0.00	(0.00, 0.00)			2	0
softa	dayssince-lastvisit	1.50	1.73	3.00	(0.00, 3.00)	0.00	−6.00	4	0
the-soa	dayssince-lastvisit	0.00	0.00	0.00	(0.00, 0.00)			2	0

Pro fingerprint: analog.

3 Debrief

3.1 Targets-Objekte

```
targets::tar_manifest() |>
  select(name) |>
  #print(n = Inf) |>
  knitr::kable()
```

name
config_file
config
data_files_list
data_files_dupes_excluded
data_imported

name
data_prepped
data_all_fct
test_unique_idvisit
data_long
data_wide_slim
data_separated
n_visits
course_and_uni_per_visit
data_separated_filtered
n_action_fingerprint
n_action
n_action_lt_500_fingerprint
n_action_lt_500

3.2 Pipeline-Graph

```
tar_visnetwork(targets_only = TRUE,
               outdated = TRUE)
```


```
tar_glimpse()
```

3.3 sessionInfo

```
sessioninfo::session_info()
```

```
─ Session info ─────────────────────────────────────────────────────────────
 setting  value
 version  R version 4.5.1 (2025-06-13)
 os       Ubuntu 25.10
 system   x86_64, linux-gnu
 ui       X11
 language (EN)
 collate  de_DE.UTF-8
 ctype    de_DE.UTF-8
 tz       Europe/Berlin
 date     2025-12-05
 pandoc   3.6.3 @ /snap/rstudio/25/resources/app/bin/quarto/bin/tools/x86_64/
(via rmarkdown)
 quarto   1.7.32 @ /snap/rstudio/25/resources/app/bin/quarto/bin/quarto

─ Packages ─────────────────────────────────────────────────────────────────
 package          * version    date (UTC) lib source
 backports          1.5.0      2024-05-23 [3] CRAN (R 4.4.1)
 base64url          1.4        2018-05-14 [3] CRAN (R 4.0.1)
 BayesFactor        0.9.12-4.7 2024-01-24 [3] CRAN (R 4.3.2)
 bayestestR       * 0.17.0     2025-08-29 [1] RSPM (R 4.5.1)
```

```
boot               1.3-31     2024-08-28 [4] CRAN (R 4.4.1)
callr              3.7.6      2024-03-25 [3] CRAN (R 4.4.0)
chromote           0.5.1      2025-04-24 [1] CRAN (R 4.5.1)
cli                3.6.5      2025-04-23 [1] CRAN (R 4.5.1)
coda               0.19-4.1   2024-01-31 [1] RSPM
codetools          0.2-20     2024-03-31 [4] CRAN (R 4.3.3)
correlation      * 0.8.8      2025-07-08 [1] RSPM (R 4.5.1)
data.table         1.17.8     2025-07-10 [1] RSPM (R 4.5.1)
datawizard       * 1.3.0      2025-10-11 [1] RSPM (R 4.5.1)
dichromat          2.0-0.1    2022-05-02 [3] CRAN (R 4.2.0)
digest             0.6.39     2025-11-19 [1] CRAN (R 4.5.1)
dplyr            * 1.1.4      2023-11-17 [3] CRAN (R 4.4.2)
easystats        * 0.7.5      2025-07-11 [1] RSPM (R 4.5.1)
effectsize       * 1.0.1      2025-05-27 [1] RSPM (R 4.5.1)
emmeans            1.10.7     2025-01-31 [3] CRAN (R 4.4.2)
estimability       1.5.1      2024-05-12 [3] CRAN (R 4.4.2)
evaluate           1.0.5      2025-08-27 [1] CRAN (R 4.5.1)
farver             2.1.2      2024-05-13 [3] CRAN (R 4.4.1)
fastmap            1.2.0      2024-05-15 [3] CRAN (R 4.4.1)
forcats          * 1.0.0      2023-01-29 [3] CRAN (R 4.2.2)
fs                 1.6.6      2025-04-12 [1] CRAN (R 4.5.1)
generics           0.1.4      2025-05-09 [1] CRAN (R 4.5.1)
ggplot2          * 4.0.1      2025-11-14 [1] RSPM (R 4.5.1)
ggstatsplot      * 0.13.3     2025-10-05 [1] RSPM
glue               1.8.0      2024-09-30 [3] CRAN (R 4.4.2)
gt               * 1.1.0      2025-09-23 [1] RSPM (R 4.5.1)
gtable             0.3.6      2024-10-25 [3] CRAN (R 4.4.2)
hms                1.1.3      2023-03-21 [3] CRAN (R 4.3.1)
htmltools          0.5.8.1    2024-04-04 [3] CRAN (R 4.4.0)
htmlwidgets        1.6.4      2023-12-06 [3] CRAN (R 4.3.2)
igraph             2.1.4      2025-01-23 [3] CRAN (R 4.5.0)
insight          * 1.4.2      2025-09-02 [1] RSPM (R 4.5.1)
jsonlite           2.0.0      2025-03-27 [1] CRAN (R 4.5.1)
knitr              1.50       2025-03-16 [3] CRAN (R 4.4.3)
labeling           0.4.3      2023-08-29 [3] CRAN (R 4.3.1)
later              1.4.4      2025-08-27 [1] RSPM (R 4.5.1)
lattice            0.22-7     2025-04-02 [4] CRAN (R 4.4.3)
lifecycle          1.0.4      2023-11-07 [3] CRAN (R 4.3.2)
lubridate        * 1.9.4      2024-12-08 [3] CRAN (R 4.4.2)
magrittr           2.0.4      2025-09-12 [1] CRAN (R 4.5.1)
MASS               7.3-65     2025-02-28 [4] CRAN (R 4.4.3)
Matrix             1.7-3      2025-03-11 [4] CRAN (R 4.4.3)
MatrixModels       0.5-4      2025-03-26 [3] CRAN (R 4.4.3)
modelbased       * 0.13.0     2025-08-30 [1] RSPM (R 4.5.1)
multcomp           1.4-28     2025-01-29 [3] CRAN (R 4.4.2)
mvtnorm            1.3-3      2025-01-10 [1] RSPM
otel               0.2.0      2025-08-29 [1] RSPM (R 4.5.1)
paletteer          1.6.0      2024-01-21 [1] RSPM
```

```
parameters       * 0.28.2     2025-09-10 [1] RSPM (R 4.5.1)
patchwork          1.3.2      2025-08-25 [1] RSPM (R 4.5.1)
pbapply            1.7-2      2023-06-27 [3] CRAN (R 4.3.1)
performance      * 0.15.2     2025-10-06 [1] RSPM (R 4.5.1)
pillar             1.11.1     2025-09-17 [1] CRAN (R 4.5.1)
pkgconfig          2.0.3      2019-09-22 [3] CRAN (R 4.0.1)
prettyunits        1.2.0      2023-09-24 [3] CRAN (R 4.3.1)
processx           3.8.6      2025-02-21 [3] CRAN (R 4.4.3)
promises           1.5.0      2025-11-01 [1] RSPM (R 4.5.1)
ps                 1.9.0      2025-02-18 [3] CRAN (R 4.4.3)
purrr            * 1.2.0      2025-11-04 [1] RSPM (R 4.5.1)
quarto             1.5.1      2025-09-04 [1] CRAN (R 4.5.1)
R6                 2.6.1      2025-02-15 [3] CRAN (R 4.4.3)
RColorBrewer       1.1-3      2022-04-03 [3] CRAN (R 4.2.0)
Rcpp               1.1.0      2025-07-02 [3] CRAN (R 4.5.1)
RcppParallel       5.1.7      2023-02-27 [3] CRAN (R 4.5.0)
readr            * 2.1.6      2025-11-14 [1] RSPM
rematch2           2.1.2      2020-05-01 [3] CRAN (R 4.0.1)
report           * 0.6.2      2025-11-03 [1] RSPM (R 4.5.1)
rlang              1.1.6      2025-04-11 [1] CRAN (R 4.5.1)
rmarkdown          2.30       2025-09-28 [1] RSPM (R 4.5.1)
rstantools         2.4.0      2024-01-31 [3] CRAN (R 4.3.2)
rstudioapi         0.17.1     2024-10-22 [3] CRAN (R 4.4.1)
S7                 0.2.1      2025-11-14 [1] RSPM (R 4.5.1)
sandwich           3.1-1      2024-09-15 [3] CRAN (R 4.4.1)
sass               0.4.10     2025-04-11 [1] RSPM (R 4.5.1)
scales           * 1.4.0      2025-04-24 [1] RSPM (R 4.5.1)
secretbase         1.0.5      2025-03-04 [1] RSPM
see              * 0.12.0     2025-09-14 [1] RSPM (R 4.5.1)
sessioninfo        1.2.3      2025-02-05 [3] CRAN (R 4.4.3)
statsExpressions   1.7.1      2025-07-27 [1] RSPM
stringi            1.8.7      2025-03-27 [1] CRAN (R 4.5.1)
stringr          * 1.6.0      2025-11-04 [1] CRAN (R 4.5.1)
survival           3.8-3      2024-12-17 [4] CRAN (R 4.4.2)
targets          * 1.11.4     2025-09-13 [1] RSPM
TH.data            1.1-3      2025-01-17 [3] CRAN (R 4.4.2)
tibble           * 3.3.0      2025-06-08 [1] CRAN (R 4.5.1)
tidyr            * 1.3.1      2024-01-24 [3] CRAN (R 4.3.2)
tidyselect         1.2.1      2024-03-11 [3] CRAN (R 4.4.0)
tidyverse        * 2.0.0      2023-02-22 [3] CRAN (R 4.4.2)
timechange         0.3.0      2024-01-18 [3] CRAN (R 4.4.3)
tinytable          0.15.1     2025-11-02 [1] CRAN (R 4.5.1)
tzdb               0.5.0      2025-03-15 [3] CRAN (R 4.4.3)
vctrs              0.6.5      2023-12-01 [3] CRAN (R 4.3.2)
visdat           * 0.6.0      2023-02-02 [1] CRAN (R 4.5.1)
visNetwork         2.1.2      2022-09-29 [3] CRAN (R 4.4.1)
webshot            0.5.5      2023-06-26 [3] CRAN (R 4.3.1)
webshot2           0.1.2      2025-04-23 [1] CRAN (R 4.5.1)
```

```
websocket          1.4.4      2025-04-10 [1] CRAN (R 4.5.1)
withr              3.0.2      2024-10-28 [3] CRAN (R 4.4.1)
xfun               0.54       2025-10-30 [1] CRAN (R 4.5.1)
xml2               1.5.0      2025-11-17 [1] CRAN (R 4.5.1)
xtable             1.8-4      2019-04-21 [3] CRAN (R 4.0.1)
yaml               2.3.10     2024-07-26 [3] CRAN (R 4.4.1)
zeallot            0.2.0      2025-05-27 [1] RSPM
zoo                1.8-14     2025-04-10 [3] CRAN (R 4.4.3)

[1] /home/sebastian-sauer/R/x86_64-pc-linux-gnu-library/4.5
[2] /usr/local/lib/R/site-library
[3] /usr/lib/R/site-library
[4] /usr/lib/R/library
* ── Packages attached to the search path.

──────────────────────────────────────────────────────────────────────────────
```