R tables easy to make?

RUG LSHTM

21 August, 2020

Table of Contents

[Table 1a: simple one-way table with default *tableby* options 💩 3](#_Toc48869557)

[Table 1a(ii): formatted 😄 4](#_Toc48869558)

[Table 1b: add/modify human readable variable labels 👍👍 5](#_Toc48869559)

[Table 1c: change default statistics and formatting 👍👍👍 6](#_Toc48869560)

[Table 1d: more than one group 💪✊` 7](#_Toc48869561)

[Table 1e: change default statistics, more than one group 👍👍 9](#_Toc48869562)

[Table 1f: Set global table options and create dataframe 👍 11](#_Toc48869563)

[Table 2: Odds ratio table 12](#_Toc48869564)

# Table 1a: simple one-way table with default *tableby* options 💩

table1a\_i <- tableby(~ age + age.factor + sex.factor + obstruct.factor + nodes,  
 data = colon\_s)  
  
summary(table1a\_i, text = T)

##   
##   
## | | Overall (N=929) |  
## |:--------------|:---------------:|  
## |Age (years) | |  
## |- Mean (SD) | 59.755 (11.949) |  
## |- Range | 18.000 - 85.000 |  
## |Age | |  
## |- <40 years | 70 (7.5%) |  
## |- 40-59 years | 344 (37.0%) |  
## |- 60+ years | 515 (55.4%) |  
## |Sex | |  
## |- Female | 445 (47.9%) |  
## |- Male | 484 (52.1%) |  
## |Obstruction | |  
## |- N-Miss | 21 |  
## |- No | 732 (80.6%) |  
## |- Yes | 176 (19.4%) |  
## |nodes | |  
## |- N-Miss | 18 |  
## |- Mean (SD) | 3.660 (3.573) |  
## |- Range | 0.000 - 33.000 |

## Table 1a(ii): formatted 😄

summary(table1a\_i, text = T) %>%  
 kable(format = "pandoc",   
 caption = "Formatted table",   
 padding = 0,   
 label = "",   
 align = c("l" , "c"),   
 col\_order = order  
 )

Formatted table

|  |  |
| --- | --- |
|  | Overall (N=929) |
| Age (years) |  |
| - Mean (SD) | 59.755 (11.949) |
| - Range | 18.000 - 85.000 |
| Age |  |
| - <40 years | 70 (7.5%) |
| - 40-59 years | 344 (37.0%) |
| - 60+ years | 515 (55.4%) |
| Sex |  |
| - Female | 445 (47.9%) |
| - Male | 484 (52.1%) |
| Obstruction |  |
| - N-Miss | 21 |
| - No | 732 (80.6%) |
| - Yes | 176 (19.4%) |
| nodes |  |
| - N-Miss | 18 |
| - Mean (SD) | 3.660 (3.573) |
| - Range | 0.000 - 33.000 |

## Table 1b: add/modify human readable variable labels 👍👍

age sex.factor rx.factor   
 "Age (years)" "Sex" "Treatment"   
 obstruct.factor perfor.factor adhere.factor   
 "Obstruction" "Perforation" "Adherence"   
 differ.factor extent.factor surg.factor   
"Differentiation" "Extent of spread" "Time from surgery"   
 node4.factor status.factor age.factor

“>4 positive nodes” “Status” “Age” loccomp.factor time.years mort\_5yr “Local complications” “Time (years)” “Mortality 5 year”

table1b <- tableby(~ age + age.factor + sex.factor + obstruct.factor + nodes,  
 data = colon\_s)  
  
summary(table1b, text = T) %>%  
 kable(format = "pandoc",   
 caption = "Formatted table",   
 padding = 0,   
 label = "",   
 align = c("l" , "c"),   
 col\_order = order  
 )

Formatted table

|  |  |
| --- | --- |
|  | Overall (N=929) |
| Age (years) |  |
| - Mean (SD) | 59.755 (11.949) |
| - Range | 18.000 - 85.000 |
| Age group |  |
| - 60+ years | 515 (55.4%) |
| - 40-59 years | 344 (37.0%) |
| - <40 years | 70 (7.5%) |
| Gender |  |
| - Female | 445 (47.9%) |
| - Male | 484 (52.1%) |
| Obstruction |  |
| - N-Miss | 21 |
| - No | 732 (80.6%) |
| - Yes | 176 (19.4%) |
| nodes |  |
| - N-Miss | 18 |
| - Mean (SD) | 3.660 (3.573) |
| - Range | 0.000 - 33.000 |

## Table 1c: change default statistics and formatting 👍👍👍

tableby(~ age + age.factor + sex.factor + obstruct.factor + nodes,  
 data = colon\_s,  
 digits = 1,   
 digits.pct = 1,  
 control = tableby.control(cat.stats = c("countpct", "N", "Nmiss2"),   
 total = T, test = F,  
 numeric.stats = c("meansd", "medianq1q3", "range", "N", "Nmiss2"),   
 stats.labels = list(Nmiss2 = "Number missing")  
 )  
 ) %>%  
 summary(text = T) %>%  
 kable(format = "pandoc",   
 caption = "Formatted table",   
 padding = 0,   
 label = "",   
 align = c("l" , "c"),   
 col\_order = order  
 )

Formatted table

|  |  |
| --- | --- |
|  | Overall (N=929) |
| Age (years) |  |
| - meansd | 59.8 (11.9) |
| - medianq1q3 | 61.0 (53.0, 69.0) |
| - range | 18.0 - 85.0 |
| - N | 929 |
| - Number missing | 0 |
| Age group |  |
| - 60+ years | 515 (55.4%) |
| - 40-59 years | 344 (37.0%) |
| - <40 years | 70 (7.5%) |
| - N | 929 |
| - Number missing | 0 |
| Gender |  |
| - Female | 445 (47.9%) |
| - Male | 484 (52.1%) |
| - N | 929 |
| - Number missing | 0 |
| Obstruction |  |
| - No | 732 (80.6%) |
| - Yes | 176 (19.4%) |
| - N | 908 |
| - Number missing | 21 |
| nodes |  |
| - meansd | 3.7 (3.6) |
| - medianq1q3 | 2.0 (1.0, 5.0) |
| - range | 0.0 - 33.0 |
| - N | 911 |
| - Number missing | 18 |

## Table 1d: more than one group 💪✊`

tableby(mort\_5yr ~ age + age.factor + sex.factor + obstruct.factor + nodes + perfor.factor,  
 data = colon\_s,  
 digits = 1,   
 digits.pct = 1,  
 control = tableby.control(cat.stats = c("countrowpct", "N", "Nmiss2"), # now with row percentages  
 total = T, test = T,  
 numeric.stats = c("meansd", "medianq1q3", "range", "N", "Nmiss2"),   
 stats.labels = list(Nmiss2 = "Number missing")  
 )  
 ) %>%  
 summary(text = T) %>%  
 kable(format = "pandoc",   
 caption = "Formatted table",   
 padding = 0,   
 label = "",   
 align = c("l" , "c"),   
 col\_order = order  
 )

Formatted table

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Alive (N=511) | Died (N=404) | Total (N=915) | p value |
| Age (years) |  |  |  | 0.986 |
| - meansd | 59.8 (11.4) | 59.9 (12.5) | 59.8 (11.9) |  |
| - medianq1q3 | 60.0 (53.0, 68.0) | 62.0 (52.0, 69.0) | 61.0 (53.0, 69.0) |  |
| - range | 22.0 - 83.0 | 18.0 - 85.0 | 18.0 - 85.0 |  |
| - N | 511 | 404 | 915 |  |
| - Number missing | 0 | 0 | 0 |  |
| Age group |  |  |  | 0.020 |
| - 60+ years | 272 (53.4%) | 237 (46.6%) | 509 (100.0%) |  |
| - 40-59 years | 208 (61.4%) | 131 (38.6%) | 339 (100.0%) |  |
| - <40 years | 31 (46.3%) | 36 (53.7%) | 67 (100.0%) |  |
| - N | 511 | 404 | 915 |  |
| - Number missing | 0 | 0 | 0 |  |
| Gender |  |  |  | 0.889 |
| - Female | 243 (55.6%) | 194 (44.4%) | 437 (100.0%) |  |
| - Male | 268 (56.1%) | 210 (43.9%) | 478 (100.0%) |  |
| - N | 511 | 404 | 915 |  |
| - Number missing | 0 | 0 | 0 |  |
| Obstruction |  |  |  | 0.189 |
| - No | 408 (56.7%) | 312 (43.3%) | 720 (100.0%) |  |
| - Yes | 89 (51.1%) | 85 (48.9%) | 174 (100.0%) |  |
| - N | 497 | 397 | 894 |  |
| - Number missing | 14 | 7 | 21 |  |
| nodes |  |  |  | < 0.001 |
| - meansd | 2.7 (2.4) | 4.9 (4.4) | 3.7 (3.6) |  |
| - medianq1q3 | 2.0 (1.0, 3.0) | 4.0 (2.0, 7.0) | 2.0 (1.0, 5.0) |  |
| - range | 1.0 - 19.0 | 0.0 - 33.0 | 0.0 - 33.0 |  |
| - N | 504 | 393 | 897 |  |
| - Number missing | 7 | 11 | 18 |  |
| Perforation |  |  |  | 0.671 |
| - No | 497 (56.0%) | 391 (44.0%) | 888 (100.0%) |  |
| - Yes | 14 (51.9%) | 13 (48.1%) | 27 (100.0%) |  |
| - N | 511 | 404 | 915 |  |
| - Number missing | 0 | 0 | 0 |  |

## Table 1e: change default statistics, more than one group 👍👍

tableby(mort\_5yr ~ age + age.factor + sex.factor + obstruct.factor + nodes + perfor.factor,  
 data = colon\_s,  
 digits = 1,   
 digits.pct = 1,  
 control = tableby.control(cat.stats = c("countrowpct", "N", "Nmiss2"), # now with row percentages  
 numeric.stats = c("meansd", "medianq1q3", "range", "N", "Nmiss2"),   
 stats.labels = list(Nmiss2 = "Number miss"),  
 total = T, test = F  
 )  
 ) %>%  
 summary(text = T) %>%  
 kable(format = "pandoc",   
 caption = "Formatted table",   
 padding = 0,   
 label = "",   
 align = c("l" , "c"),   
 col\_order = order  
 )

Formatted table

|  |  |  |  |
| --- | --- | --- | --- |
|  | Alive (N=511) | Died (N=404) | Total (N=915) |
| Age (years) |  |  |  |
| - meansd | 59.8 (11.4) | 59.9 (12.5) | 59.8 (11.9) |
| - medianq1q3 | 60.0 (53.0, 68.0) | 62.0 (52.0, 69.0) | 61.0 (53.0, 69.0) |
| - range | 22.0 - 83.0 | 18.0 - 85.0 | 18.0 - 85.0 |
| - N | 511 | 404 | 915 |
| - Number miss | 0 | 0 | 0 |
| Age group |  |  |  |
| - 60+ years | 272 (53.4%) | 237 (46.6%) | 509 (100.0%) |
| - 40-59 years | 208 (61.4%) | 131 (38.6%) | 339 (100.0%) |
| - <40 years | 31 (46.3%) | 36 (53.7%) | 67 (100.0%) |
| - N | 511 | 404 | 915 |
| - Number miss | 0 | 0 | 0 |
| Gender |  |  |  |
| - Female | 243 (55.6%) | 194 (44.4%) | 437 (100.0%) |
| - Male | 268 (56.1%) | 210 (43.9%) | 478 (100.0%) |
| - N | 511 | 404 | 915 |
| - Number miss | 0 | 0 | 0 |
| Obstruction |  |  |  |
| - No | 408 (56.7%) | 312 (43.3%) | 720 (100.0%) |
| - Yes | 89 (51.1%) | 85 (48.9%) | 174 (100.0%) |
| - N | 497 | 397 | 894 |
| - Number miss | 14 | 7 | 21 |
| nodes |  |  |  |
| - meansd | 2.7 (2.4) | 4.9 (4.4) | 3.7 (3.6) |
| - medianq1q3 | 2.0 (1.0, 3.0) | 4.0 (2.0, 7.0) | 2.0 (1.0, 5.0) |
| - range | 1.0 - 19.0 | 0.0 - 33.0 | 0.0 - 33.0 |
| - N | 504 | 393 | 897 |
| - Number miss | 7 | 11 | 18 |
| Perforation |  |  |  |
| - No | 497 (56.0%) | 391 (44.0%) | 888 (100.0%) |
| - Yes | 14 (51.9%) | 13 (48.1%) | 27 (100.0%) |
| - N | 511 | 404 | 915 |
| - Number miss | 0 | 0 | 0 |

## Table 1f: Set global table options and create dataframe 👍

# Set global table options  
  
mycontrols <- tableby.control(test=FALSE, total=FALSE,  
 numeric.test="kwt", cat.test="chisq",  
 numeric.stats=c("meansd", "medianq1q3", "range", "N", "Nmiss2"),  
 cat.stats=c("countpct", "N", "Nmiss2"),  
 stats.labels=list(N = 'Count',  
 median = 'Median',   
 q1q3 = 'Q1,Q3'),  
 digits = 1,  
 digits.pct = 1  
 )  
  
# Create a table as data frame  
  
table\_1f <- tableby(mort\_5yr ~ age + age.factor + sex.factor,  
 data = colon\_s,  
 control = mycontrols) %>%  
 summary(text = NULL) # set to null to remove any text formatting   
  
print(as.data.frame(table\_1f))

## Alive (N=511) Died (N=404)  
## 1 Age (years)   
## 2 meansd 59.8 (11.4) 59.9 (12.5)  
## 3 medianq1q3 60.0 (53.0, 68.0) 62.0 (52.0, 69.0)  
## 4 range 22.0 - 83.0 18.0 - 85.0  
## 5 Count 511 404  
## 6 Nmiss2 0 0  
## 7 Age group   
## 8 60+ years 272 (53.2%) 237 (58.7%)  
## 9 40-59 years 208 (40.7%) 131 (32.4%)  
## 10 <40 years 31 (6.1%) 36 (8.9%)  
## 11 Count 511 404  
## 12 Nmiss2 0 0  
## 13 Gender   
## 14 Female 243 (47.6%) 194 (48.0%)  
## 15 Male 268 (52.4%) 210 (52.0%)  
## 16 Count 511 404  
## 17 Nmiss2 0 0

*Other possibilities* + Subsets of groups (e.g. pregnant women) \* Weighted estimates \* Apply custom statistical tests for certain variables \* Additional groupings by strata \* Full list here: <https://cran.r-project.org/web/packages/arsenal/vignettes/tableby.html>

# Table 2: Odds ratio table

model\_output1 %>%  
 kable(format = "pandoc",   
 caption = "",   
 padding = 0,   
 label = "",   
 align = c("l" , "c"),   
 col\_order = order,  
 row.names = NA,  
 col.names = NA  
 )

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | Dependent: Mortality 5 year |  | Alive | Died | OR (univariable) | OR (multivariable) |
| 3 | Age group | 60+ years | 272 (53.4) | 237 (46.6) | - | - |
| 2 |  | 40-59 years | 208 (61.4) | 131 (38.6) | 0.72 (0.55-0.96, p=0.023) | 0.71 (0.53-0.94, p=0.017) |
| 1 |  | <40 years | 31 (46.3) | 36 (53.7) | 1.33 (0.80-2.23, p=0.270) | 1.23 (0.74-2.08, p=0.426) |
| 8 | Gender | Female | 243 (55.6) | 194 (44.4) | - | - |
| 9 |  | Male | 268 (56.1) | 210 (43.9) | 0.98 (0.76-1.27, p=0.889) | 0.98 (0.75-1.28, p=0.902) |
| 4 | Obstruction | No | 408 (56.7) | 312 (43.3) | - | - |
| 5 |  | Yes | 89 (51.1) | 85 (48.9) | 1.25 (0.90-1.74, p=0.189) | 1.25 (0.90-1.76, p=0.186) |
| 6 | Perforation | No | 497 (56.0) | 391 (44.0) | - | - |
| 7 |  | Yes | 14 (51.9) | 13 (48.1) | 1.18 (0.54-2.55, p=0.672) | 1.12 (0.51-2.44, p=0.770) |

*Lots of features and customisations* \* Full list here: <https://finalfit.org/articles/all_tables_examples.html>