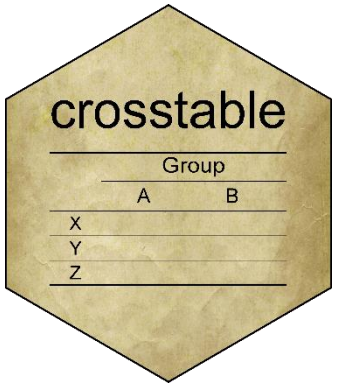
#RR2023 Rencontres R AVIGNON 2023



https://github.com/DanChaltiel/crosstable

Crosstable

L'analyse descriptive pour les nuls

(et les autres)

GUSTAVE ROUSSY
CANCER CAMPUS GRAND PARIS

Dan Chaltiel

Biostatisticien

Institut Gustave Roussy

cesp

Oncostat

Centre de recherche en épidémiologie et santé des populations

Crosstable: exemple 1

```
library(crosstable)
ct1 = crosstable(mtcars2, c(disp, vs), by=am, total="both",
                 percent_pattern="{n} ({p_row}/{p_col})", percent_digits=0) %>%
  as_flextable()
ct1
```

by=am

total="both" (row)

label	variable	Transmission		Total
		auto	manual	
Displacement (cu.in.)	Min / Max	120.1 / 472.0	71.1 / 351.0	71.1 / 472.0
	Med [IQR]	275.8 [196.3;360.0]	120.3 [79.0;160.0]	196.3 [120.8;326.0]
	Mean (std)	290.4 (110.2)	143.5 (87.2)	230.7 (123.9)
	N (NA)	19 (0)	13 (0)	32 (0)
Engine	straight	7 (50%/37%)	7 (50%/54%)	14 (44%)
	vshaped	12 (67%/63%)	6 (33%/46%)	18 (56%)
	Total	19 (59%)	13 (41%)	32 (100%)

c(disp, vs)

total="both" (col)

percent_pattern="{n} ({p_row}/{p_col})«
percent_digits=0

Crosstable: exemple 1

```
library(crosstable)
ct1 = crosstable(mtcars2, c(disp, vs), by=am, total="both",
                 percent_pattern="{n} ({p_row}/{p_col})", percent_digits=0) %>%
  as_flextable()
ct1
```

by=am

total="both" (row)

label	variable	Transmission		Total
		auto	manual	
Displacement (cu.in.)	Min / Max	120.1 / 472.0	71.1 / 351.0	71.1 / 472.0
	Med [IQR]	275.8 [196.3;360.0]	120.3 [79.0;160.0]	196.3 [120.8;326.0]
	Mean (std)	290.4 (110.2)	143.5 (87.2)	230.7 (123.9)
	N (NA)	19 (0)	13 (0)	32 (0)
Engine	straight	7 (50%/37%)	7 (50%/54%)	14 (44%)
	vshaped	12 (67%/63%)	6 (33%/46%)	18 (56%)
	Total	19 (59%)	13 (41%)	32 (100%)

c(disp, vs)

total="both" (col)

percent_pattern="{n} ({p_row}/{p_col})«
percent_digits=0

Crosstable: exemple 2

```
ct2 = crosstable(mtcars2, c(starts_with("cy"), ends_with("at")), by=c(am, vs),
                 label=FALSE, num_digits=3, funs=c(mean, quantile),
                 funs_arg=list(probs=c(.25,.75))) %>%
  as_flextable(compact=TRUE, header_show_n=1:2)
ct2
```

header_show_n=1:2
Shows the group size

by=c(am, vs)

variable	vs=straight (n=14)		vs=vshaped (n=18)	
	am=auto (n=7)	am=manual (n=7)	am=auto (n=12)	am=manual (n=6)
cyl				
4	3 (27.27%)	7 (63.64%)	0 (0%)	1 (9.09%)
6	4 (57.14%)	0 (0%)	0 (0%)	3 (42.86%)
8	0 (0%)	0 (0%)	12 (85.71%)	2 (14.29%)
drat				
mean	3.570	4.149	3.121	3.935
quantile 25%	3.385	3.965	3.052	3.690
quantile 75%	3.920	4.165	3.165	4.140

starts_with("cy")

ends_with("at")

funs=c(mean, quantile)

funs_arg=list(probs=c(.25,.75))

num_digits=3

Crosstable: exemple 2

```
ct2 = crosstable(mtcars2, c(starts_with("cy"), ends_with("at")), by=c(am, vs),
                 label=FALSE, num_digits=3, funs=c(mean, quantile),
                 funs_arg=list(probs=c(.25,.75))) %>%
  as_flextable(compact=TRUE, header_show_n=1:2)
ct2
```

header_show_n=1:2
Shows the group size

by=c(am, vs)

variable	vs=straight (n=14)		vs=vshaped (n=18)	
	am=auto (n=7)	am=manual (n=7)	am=auto (n=12)	am=manual (n=6)
cyl				
4	3 (27.27%)	7 (63.64%)	0 (0%)	1 (9.09%)
6	4 (57.14%)	0 (0%)	0 (0%)	3 (42.86%)
8	0 (0%)	0 (0%)	12 (85.71%)	2 (14.29%)
drat				
mean	3.570	4.149	3.121	3.935
quantile 25%	3.385	3.965	3.052	3.690
quantile 75%	3.920	4.165	3.165	4.140

starts_with("cy")

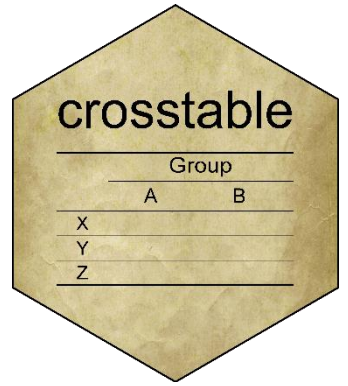
ends_with("at")

funs=c(mean, quantile)

funs_arg=list(probs=c(.25,.75))

num_digits=3

Crosstable: résumé



- Alternative à {table1} et {gtsummary}
- Tidyverse FTW!
  - Tidyselection sur les colonnes
  - Lambda-fonctions pour funs
- Tout est paramétrable via crosstable_options()
- Tout est sur github:
  - Code : https://github.com/DanChaltiel/crosstable
  - Doc : https://danchaltiel.github.io/crosstable/reference/crosstable.html

Usage

```
crosstable(
  data,
  cols = everything(),
  ...,
  by = NULL,
  total = c("none", "row", "column", "both"),
  percent_pattern = "{n} ({p_row})",
  percent_digits = 2,
  num_digits = 1,
  showNA = c("ifany", "always", "no"),
  label = TRUE,
  funs = c(` ` = cross_summary),
  funs_arg = list(),
  cor_method = c("pearson", "kendall", "spearman"),
  drop_levels = FALSE,
  unique_numeric = 3,
  date_format = NULL,
  times = NULL,
  followup = FALSE,
  test = FALSE,
  test_args = crosstable_test_args(),
  effect = FALSE,
  effect_args = crosstable_effect_args(),
  margin = deprecated(),
  .vars = deprecated()
)
```

Crosstable: rapports

- Interface avec {officer} pour l'édition de rapports automatiques

```
1  library(crosstable)
2  library(officer)
3
4  ct1 = crosstable(iris2, by=Species)
5
6  crosstable_options(
7    crosstable_fontsize_body=8,
8    crosstable_padding_v=0,
9  )
10
11 doc = read_docx() %>%
12   body_add_title("Dataset iris (nrow={nrow(iris2)})", 1) %>%
13   body_add_title("Not compacted", 2) %>%
14   body_add_normal("Table @ref(ct1) is an *example*.") %>%
15   body_add_crosstable(ct1) %>%
16   body_add_table_legend("Example table",
17                         bookmark="ct1")
18
19 write_and_open(doc)
```

1. Dataset iris (nrow=150)

1.1. Not compacted

Table **1** is an *example*.

label	variable	Specie		
		setosa	versicolor	virginica
Length of Sepal	Min / Max	4.3 / 5.8	4.9 / 7.0	4.9 / 7.9
	Med [IQR]	5.0 [4.8;5.2]	5.9 [5.6;6.3]	6.5 [6.2;6.9]
	Mean (std)	5.0 (0.4)	5.9 (0.5)	6.6 (0.6)
	N (NA)	50 (0)	50 (0)	50 (0)
Width of Sepal	Min / Max	2.3 / 4.4	2.0 / 3.4	2.2 / 3.8
	Med [IQR]	3.4 [3.2;3.7]	2.8 [2.5;3.0]	3.0 [2.8;3.2]
	Mean (std)	3.4 (0.4)	2.8 (0.3)	3.0 (0.3)
	N (NA)	50 (0)	50 (0)	50 (0)
Length of Petal	Min / Max	1.0 / 1.9	3.0 / 5.1	4.5 / 6.9
	Med [IQR]	1.5 [1.4;1.6]	4.3 [4.0;4.6]	5.5 [5.1;5.9]
	Mean (std)	1.5 (0.2)	4.3 (0.5)	5.6 (0.6)
	N (NA)	50 (0)	50 (0)	50 (0)
Width of Petal	Min / Max	0.1 / 0.6	1.0 / 1.8	1.4 / 2.5
	Med [IQR]	0.2 [0.2;0.3]	1.3 [1.2;1.5]	2.0 [1.8;2.3]
	Mean (std)	0.2 (0.1)	1.3 (0.2)	2.0 (0.3)
	N (NA)	50 (0)	50 (0)	50 (0)

Table 1: Example table

Crosstable: rapports

- Interface avec {officer} pour l'édition de rapports automatiques

```
1  library(crosstable)
2  library(officer)
3
4  ct1 = crosstable(iris2, by=Species)
5
6  crosstable_options(
7    crosstable_fontsize_body=8,
8    crosstable_padding_v=0,
9  )
10
11 doc = read_docx() %>%
12   body_add_title("Dataset iris (nrow={nrow(iris2)})", 1) %>%
13   body_add_title("Not compacted", 2) %>%
14   body_add_normal("Table @ref(ct1) is an *example*.") %>%
15   body_add_crosstable(ct1) %>%
16   body_add_table_legend("Example table",
17                         bookmark="ct1")
18
19 write_and_open(doc)
```

1. Dataset iris (nrow=150)

1.1. Not compacted

Table **1** is an *example*.

label	variable	Specie		
		setosa	versicolor	virginica
Length of Sepal	Min / Max	4.3 / 5.8	4.9 / 7.0	4.9 / 7.9
	Med [IQR]	5.0 [4.8;5.2]	5.9 [5.6;6.3]	6.5 [6.2;6.9]
	Mean (std)	5.0 (0.4)	5.9 (0.5)	6.6 (0.6)
	N (NA)	50 (0)	50 (0)	50 (0)
Width of Sepal	Min / Max	2.3 / 4.4	2.0 / 3.4	2.2 / 3.8
	Med [IQR]	3.4 [3.2;3.7]	2.8 [2.5;3.0]	3.0 [2.8;3.2]
	Mean (std)	3.4 (0.4)	2.8 (0.3)	3.0 (0.3)
	N (NA)	50 (0)	50 (0)	50 (0)
Length of Petal	Min / Max	1.0 / 1.9	3.0 / 5.1	4.5 / 6.9
	Med [IQR]	1.5 [1.4;1.6]	4.3 [4.0;4.6]	5.5 [5.1;5.9]
	Mean (std)	1.5 (0.2)	4.3 (0.5)	5.6 (0.6)
	N (NA)	50 (0)	50 (0)	50 (0)
Width of Petal	Min / Max	0.1 / 0.6	1.0 / 1.8	1.4 / 2.5
	Med [IQR]	0.2 [0.2;0.3]	1.3 [1.2;1.5]	2.0 [1.8;2.3]
	Mean (std)	0.2 (0.1)	1.3 (0.2)	2.0 (0.3)
	N (NA)	50 (0)	50 (0)	50 (0)

Table 1: Example table

Crosstable: pivot/transpose

Pivot : `pivot_wider()` sur les niveaux

```
ct1 = crosstable(mtcars2, c(mpg, drat, wt, qsec), label=F,
                 funs=c(mean, sd))
```

```
ct1 %>% as_flextable()
```

label	variable	value
mpg	mean	20.1
	sd	6.0
drat	mean	3.6
	sd	0.5
wt	mean	3.2
	sd	1.0
qsec	mean	17.8
	sd	1.8

```
pivot_crosstable(ct1) %>%
  as_flextable()
```

→

label	variable	
	mean	sd
mpg	20.1	6.0
drat	3.6	0.5
wt	3.2	1.0
qsec	17.8	1.8

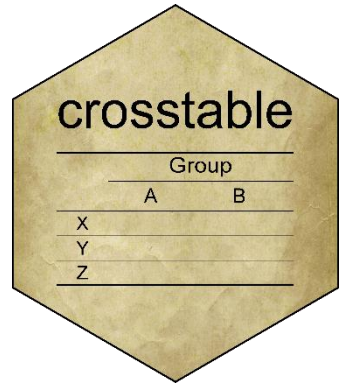
Transpose : inverse `cols` & by

```
ct2 = crosstable(mtcars2, c(mpg, drat), by=am, label=F)
```

```
transpose_crosstable(ct2) %>%
  as_flextable()
```

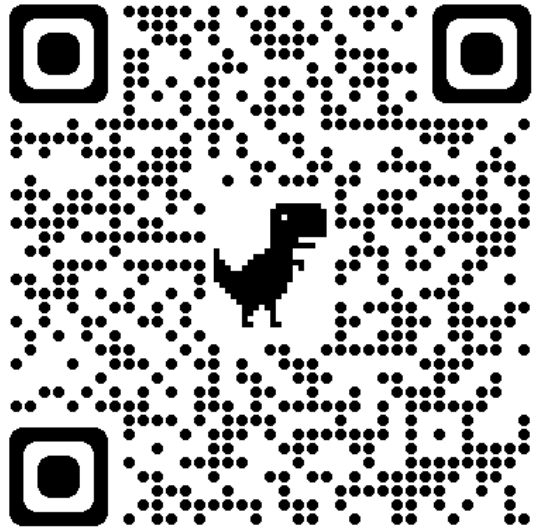
am	variable	label	
		mpg	drat
auto	Min / Max	10.4 / 24.4	2.8 / 3.9
	Med [IQR]	17.3 [14.9;19.2]	3.1 [3.1;3.7]
	Mean (std)	17.1 (3.8)	3.3 (0.4)
	N (NA)	19 (0)	19 (0)
manual	Min / Max	15.0 / 33.9	3.5 / 4.9
	Med [IQR]	22.8 [21.0;30.4]	4.1 [3.8;4.2]
	Mean (std)	24.4 (6.2)	4.0 (0.4)
	N (NA)	13 (0)	13 (0)

Crosstable: misc



- Plein d'autres fonctions 😀
- Autres formats d'output
  - `as_gt()`, `as_workbook()`
- Gestion des labels
  - `set_label()`, `apply_label()`
  - `rename_with_labels()`
- Helpers pour officer
  - `body_add_normal()`, `body_add_gg2()`
  - `generate_autofit_macro()`

Lien vers la documentation



Merci de votre attention