

**Publication-ready analytical** and summary tables with R **Cheat Sheet** 

## tbl\_summary()

Calculates descriptive stats for continuous. categorical, and dichotomous variables.

## tbl\_regression()

Turns a regression model object into a customized, formatted table.

**Core Table Functions** 

#### tbl survfit()

Turns a survfit object into a customized table with time-to-event estimates.



tbl\_summary() using tidyverse syntax to summarize specific columns of a dataset with flexible customization options (See vignette!)

#### Basic code

trial %>% select(trt,age,grade,response) %>% tbl\_summary()

#### **Basic table**

Characteristic	$N = 200^{7}$
Chemotherapy Treatment	
Drug A	98 (49%)
Drug B	102 (51%)
Age, yrs	47 (38, 57)
Unknown	11
Grade	
I	68 (34%)
II	68 (34%)
III	64 (32%)
Tumor Response	61 (32%)
Unknown	7
¹n (%); Median (IQR)	

label Formula list of variable labels

**Argument Input** 

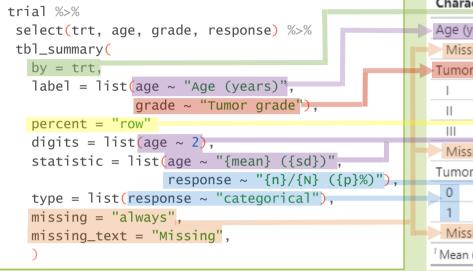
#### tbl\_svysummary() for survey objects

Same functionality as tbl summary(), but takes a survey object as input, and accounts for survey weights and design.

More info at:

http://www.danieldsjoberg.com/gtsum mary/reference/tbl\_svysummary.html

#### **Customized code**



#### **Customized table**

	Characteristic	<b>Drug A</b> , N = 98 <sup>7</sup>	<b>Drug B</b> , N = 102 <sup>7</sup>				
	Age (years)	47.01 (14.71)	47.45 (14.01)				
	Missing	7	4				
	Tumor grade						
	I	35 (51%)	33 (49%)				
	II	32 (47%)	36 (53%)				
	III	31 (48%)	33 (52%)				
H	Missing	0	0				
	Tumor Respons	e					
	0	67/132 (51%)	65/132 (49%)				
Ш	1	28/61 (46%)	33/61 (54%)				
	Missing	3	4				
	<sup>1</sup> Mean (SD); n (%); n/N (%)						

For more info on customization arguments and options, visit <a href="http://www.danieldsjoberg.com/gtsummary/reference/tbl\_summary.html">http://www.danieldsjoberg.com/gtsummary/reference/tbl\_summary.html</a>

#### **Customization options**

Changes text of variable name in table

#### Effect on table

Summary statistics will be calculated separately by Column to crosstabulate by for each level of the variable

Formula list of summary statistic Changes summary statistics displayed for statistic type for each variable specified variables in table

Formula list of number of decimal Changes number of rounded decimal places in places to display table for specified continuous variables

Formula list specifying variable Changes variable type for specified variables, affecting which summary statistics are displayed

Formula list of value to display for Changes the value displayed for dichotomous dichotomous variables type variables

Changes whether missing observations are missing "no", "ifany", "always" reported

Changes the name of the missing data level for String to display for count of missing\_text missing observations appropriate variables

Formula list of type of sorting to Changes the type of sorting for categorical sort perform ("frequency" or variables "alphanumeric")

Changes how percentage statistics are calculated percent | "column", "row", or "cell" and displayed

#### **Helper functions** useful extensions to tbl\_summary()

#### **Extended code**

add\_p()

#### trial %>% select(trt, age, response) %>% tbl\_summary( by = trt,missing= "no" ) %>% add\_n() %>% add\_overall() %>%

#### **Extended table**

Characteristic	N	<b>Overall</b> , N = 200 <sup>1</sup>	<b>Drug A</b> , N = 98 <sup>1</sup>	<b>Drug B</b> , N = 102 <sup>1</sup>	p-value
Age, yrs	189	47 (38, 57)	46 (37, 59)	48 (39, 56)	0.7
Tumor Response	193	61 (32%)	28 (29%)	33 (34%)	0.5
<sup>7</sup> Median (IQR); n (%) <sup>2</sup> Wilcoxon rank sum test; Pearson's Chi-squared test					
	+				
add n()		add c	verall()	See also:	

Adds a column with the total number of nonmissing observations

### add\_overall()

Adds a column with overall summary statistics

add\_q() bold\_p() bold\_labels() add\_stat()

#### Adds column of p-values generated by testing for differences between groups. Takes arguments below.

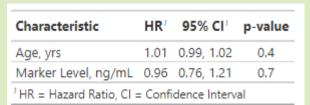
Argument	Default	Input	Effect on table
test	Continuous: "Kruskal test"; Categorical, expected cell counts ≥5: "chisq.test.no.correct"; Categorical, expected cell counts < 5: "fisher.test"	Formula list specifying statistical test to perform for each variable: "t.test", "aov", "wilcox.test", "kruskal.test", "chisq.test" and "lme4"; custom tests possible too	Changes p-value in table based on specified statistical test
pvalue_fun	style_pvalue()	Function to round and format p-values	Changes format of p-values in table
	For many info visit bttp://www.d	anialdaiahara aana/ataunanaan/rafaran	as /index basel

For more info, visit http://www.danieldsjoberg.com/gtsummary/reference/index.html

#### tbl\_regression() Present regression model object in publication-ready table

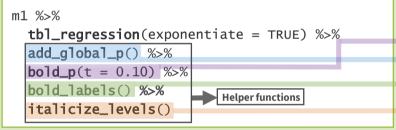
#### cox model: basic code

#### cox model: basic table



#### glm model: basic code

## glm model: custom code using helper functions



#### 1. Build model of interest

#### 2. Use tbl\_regression() to present results

tbl\_regression() supports most commonly used regression models, and uses broom::tidy(x) to perform initial tidying.

For more info:

http://www.danieldsjoberg.com/gtsummary/reference/tbl\_regression.html

#### glm model: basic table

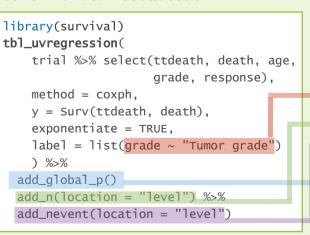
Characteristic	OR <sup>1</sup>	95% CI <sup>1</sup>	p-value		
Age, yrs	1.02	1.00, 1.04	0.091		
T Stage					
T1	_	_			
T2	0.58	0.24, 1.37	0.2		
T3	0.94	0.39, 2.28	0.9		
T4	0.79	0.33, 1.90	0.6		
OR = Odds Ratio, CI = Confidence Interval					

#### glm model: custom table

	Characteristic		OR <sup>7</sup>	95%	CI	p-value	
Age, yrs		1.02	1.00,	1.04	0.087		
	T Stage					4	0.6
	T1			_	_	-	
	T2			0.58	0.24,	1.37	
	T3			0.94	0.39,	2.28	
	T4			0.79	0.33,	1.90	
7	OR =	Odd	s Ratio	, CI = (	Confid	ence li	nterval

#### tbl\_uvregression() displays multiple univariate regression models at once

#### cox univariate models: code



#### cox univariate models: table

Characteristic	N	Event N	$\mathbf{HR}^{I}$	95% CI <sup>1</sup>	p-value		
Age, yrs	189	103	1.01	0.99, 1.02	0.3		
Tumor grade					0.075		
1	68	33	_	_			
II	68	36	1.28	0.80, 2.05			
III	64	43	1.69	1.07, 2.66			
Tumor Response	193	107	0.50	0.31, 0.78	0.001		
HR = Hazard Ratio, CI = Confidence Interval							

Requires "method" parameter specifying model type. Can estimate univariate regression models holding either outcome ("y") or covariate ("x") constant, or both (see "formula" parameter). For more info about "formula" and other parameters, see:

https://www.danieldsjoberg.com/gtsummary/reference/tbl\_uvregression.html

#### tbl\_survfit() Present survfit object with custom estimates In publication-ready table

```
library(survival)
tbl_survfit(
list(
    survfit(Surv(ttdeath, death) ~ 1, trial),
    survfit(Surv(ttdeath, death) ~ trt, trial)
),
times = c(12, 24),
label_header = "**{time} Month**"
)
```

<u>\</u>			
12 Month	24 Month		
88% (84%, 93%)	44% (38%, 51%)		
91% (85%, 97%)	47% (38%, 58%)		
86% (80%, 93%)	41% (33%, 52%)		
	88% (84%, 93%) 91% (85%, 97%)		

#### **Options for defining model:**

- 1. x=explicit survfit model (or list of models) from dataframe
- 2. x=dataframe and designate y=Surv object and include=covariates in model

#### **Options for output:**

- 1. time-specific survival estimates using times=times of interest
- 2. quantile (e.g. median) survival times using probs=quantile of interest

For more info: http://www.danieldsjoberg. com/gtsummary/reference/ tbl\_survfit.html

# library(survival) tbl\_survfit( trial, y = Surv(ttdeath, death), include = c(trt, grade), probs = 0.5, label\_header = "\*\*Median Survival\*\*" ) %>% add\_p()

Characteristic	Median Survival	p-value <sup>1</sup>
Chemotherapy Treatment		0.2
Drug A	24 (21, —)	
Drug B	21 (18, —)	
Grade		0.072
I	— (22, —)	
II	22 (18, —)	
III	20 (18, 23)	
<sup>1</sup> Log-rank test		

## tbl\_merge(), tbl\_stack() combine tables by row or column

```
t1 = tbl_survfit(
  list(survfit(Surv(ttdeath, death) ~ trt +
  grade, trial),
   times = c(12, 24),
   label_header = "**{time} Month**"
)
```

```
t2 = tbl_survfit(
   trial,
   y = Surv(ttdeath, death),
   include = c(trt, grade),
   probs = 0.5,
   label_header = "**Median Survival**"
) %>% add_p()
```

tbl\_merge(list(t1,t2), tab\_spanner = FALSE)

Characteristic	12 Month	24 Month	Median Survival	p-value <sup>1</sup>
Chemotherapy Treatment				0.2
Drug A	91% (85%, 97%)	47% (38%, 58%)	24 (21, —)	
Drug B	86% (80%, 93%)	41% (33%, 52%)	21 (18, —)	
Grade				0.072
I	97% (93%, 100%)	51% (41%, 65%)	— (22, —)	
II	82% (74%, 92%)	47% (37%, 61%)	22 (18, —)	
III	86% (78%, 95%)	33% (23%, 47%)	20 (18, 23)	
<sup>1</sup> Log-rank test				

tbl\_merge combines columns, tbl\_stack() combines rows. For more info, see https://www.danieldsjoberg.com/gtsummary/reference/