

Package ‘testflow’

July 2, 2026

Title A Workflow for Statistical Testing, Interpretation, and
'ggplot2'-Based Visualization

Version 0.8.2

Author Imad EL BADISY [aut, cre]

Maintainer Imad EL BADISY <elbadisyimad@gmail.com>

Description Provides a unified workflow for choosing, running, interpreting,
and visualizing common statistical tests. The package combines assumption
checks, test selection, effect sizes, formatted results, plain-language
interpretation, and 'ggplot2'-based statistical visualizations. Implemented
methods follow standard references including Casella and Berger (2002,
ISBN:9780534243128), Hollander et al. (2013, ISBN:9781118553299), Agresti
(2013, ISBN:9780470463635), and Cohen (1988, ISBN:9780805802832).

License MIT + file LICENSE

Encoding UTF-8

Depends R (>= 4.1.0)

RoxygenNote 7.3.3

Imports stats, cli, rlang, dplyr, tidyr, tidyselect, tibble, purrr,
broom, ggplot2

Suggests testthat (>= 3.0.0), knitr, rmarkdown, car, effectsize,
rstatix, patchwork

Config/testthat/edition 3

VignetteBuilder knitr

NeedsCompilation no

Repository CRAN

Date/Publication 2026-07-02 18:40:02 UTC

Contents

as_tibble	2
make_cardio_data	3
plot.testflow	3

print.testflow	4
report	4
report_test	5
summary.testflow	5
sumtab	6
test_categorical	7
test_correlation	8
test_correlation_matrix	9
test_factorial	10
test_groups	11
test_multinomial	12
test_one_sample	13
test_outliers	14
test_paired	14
test_paired_categorical	15
test_proportion	16
test_repeated	17
test_repeated_categorical	18
test_repeated_long	19
test_two_groups	20

Index	21
--------------	-----------

as_tibble	<i>Convert a testflow object to a one-row tibble</i>
-----------	--

Description

Convert a testflow object to a one-row tibble

Usage

```
as_tibble(x, ...)
```

Arguments

x	A testflow object.
...	Unused.

Value

A one-row tibble with the workflow name, design, variables, recommended test, null hypothesis, statistic, degrees of freedom when available, p-value, confidence interval when available, effect-size fields, and decision text.

make_cardio_data	<i>Simulate a small cardiovascular teaching dataset</i>
------------------	---

Description

Simulate a small cardiovascular teaching dataset

Usage

```
make_cardio_data(n = 180, seed = 2026)
```

Arguments

n	Number of rows.
seed	Random seed.

Value

A tibble with example numeric and categorical variables.

plot.testflow	<i>Plot a testflow object</i>
---------------	-------------------------------

Description

Plot a testflow object

Usage

```
## S3 method for class 'testflow'  
plot(x, title = NULL, subtitle = NULL, caption = NULL, ...)
```

Arguments

x	A testflow object.
title	Optional plot title override. Defaults to the stored title.
subtitle	Optional plot subtitle override. Defaults to the stored subtitle.
caption	Optional plot caption override. Defaults to the stored caption.
...	Unused.

Value

A ggplot object stored in the testflow object, with optional title, subtitle, and caption overrides applied. If the workflow was created with `plot = FALSE`, returns `NULL`.

`print.testflow` *Print a testflow object*

Description

Print a testflow object.

Usage

```
## S3 method for class 'testflow'  
print(x, ...)
```

Arguments

`x` A testflow object.
`...` Unused.

Details

Console colors are enabled by default in interactive sessions. Use `options(testflow.cli_colors = FALSE)` to disable colors, or `options(testflow.cli_colors = TRUE)` to force colors in non-interactive output.

Value

The input testflow object, invisibly. Called for its side effect of printing a formatted workflow summary to the console.

`report` *Return a ready-to-use testflow report*

Description

Return a ready-to-use testflow report

Usage

```
report(x, ...)
```

Arguments

`x` A testflow object.
`...` Unused.

Value

A length-one character vector containing a plain-language summary of the workflow result, including the design, recommended test, p-value, effect size when reported, and null hypothesis when available.

report_test	<i>Return a ready-to-use testflow report</i>
-------------	--

Description

Return a ready-to-use testflow report

Usage

```
report_test(x)
```

Arguments

x	A testflow object.
---	--------------------

Value

A length-one character vector containing the same report text as [report\(\)](#) for a testflow object.

summary.testflow	<i>Summarize a testflow object</i>
------------------	------------------------------------

Description

Summarize a testflow object.

Usage

```
## S3 method for class 'testflow'
summary(object, ...)
```

Arguments

object	A testflow object.
...	Unused.

Details

Console colors follow the same `testflow.cli_colors` option used by [print.testflow\(\)](#).

Value

A `summary.testflow` list containing the workflow metadata, descriptives, assumptions, recommended test, primary and alternative test results, post-hoc results when available, effect size, decision, and report text.

sumtab

Build a compact descriptive summary table

Description

Builds a compact table of descriptive summaries using a formula interface. Numeric variables are summarized as mean (SD); median [Q1, Q3]; n. Categorical variables are summarized as n (percent).

Usage

```
sumtab(
  formula,
  data,
  p_value = FALSE,
  overall = TRUE,
  digits = 1,
  p_digits = 3,
  alpha = 0.05,
  fisher_threshold = 5,
  na.rm = TRUE
)
```

Arguments

<code>formula</code>	A one-sided formula such as <code>~ age + sex</code> or <code>~ age + sex treatment</code> .
<code>data</code>	A data frame.
<code>p_value</code>	Logical; add a p-value column when a grouping variable is supplied.
<code>overall</code>	Logical; include an overall summary column.
<code>digits</code>	Number of digits for summary statistics.
<code>p_digits</code>	Number of digits for formatted p-values.
<code>alpha</code>	Significance level used by automatic test selection.
<code>fisher_threshold</code>	Expected-count threshold for Fisher's exact test.
<code>na.rm</code>	Logical; remove missing values before summaries and tests.

Details

When `p_value = TRUE` and a grouping variable is supplied, `sumtab()` chooses the p-value test automatically. Numeric variables use Student t-test, Welch t-test, or Wilcoxon rank-sum test for two groups, and one-way ANOVA, Welch ANOVA, or Kruskal-Wallis test for more than two groups. Categorical variables use a chi-square test unless expected counts fall below `fisher_threshold`, in which case Fisher's exact test is used.

Value

A tibble with one row per numeric variable and one row per categorical level.

Examples

```
dat <- make_cardio_data(80, seed = 1)
sumtab(~ age + sex | treatment, dat, p_value = TRUE)
```

test_categorical	<i>Test association between two categorical variables</i>
------------------	---

Description

Test association between two categorical variables

Usage

```
test_categorical(
  formula,
  data,
  y = NULL,
  alpha = 0.05,
  fisher_threshold = 5,
  plot = TRUE,
  na.rm = TRUE
)
```

Arguments

<code>formula</code>	A formula such as <code>x ~ y</code> , or a data frame when using pipe/data-first style.
<code>data</code>	A data frame, or a first categorical column when using data-first style.
<code>y</code>	Second categorical column. Optional when using formula style.
<code>alpha</code>	Significance level.
<code>fisher_threshold</code>	Expected-count threshold for Fisher's exact test.
<code>plot</code>	Logical; include a ggplot object.
<code>na.rm</code>	Logical; remove missing values.

Value

A testflow object with class `testflow_categorical`. The object is a list containing the cleaned data, categorical descriptives, assumption checks, recommended association test, primary test result with null hypothesis, alternative chi-square and Fisher results, effect size, optional `ggplot`, original call, and report text.

References

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine*, 50(302), 157-175.

Fisher, R. A. (1922). On the interpretation of chi-square from contingency tables, and the calculation of P. *Journal of the Royal Statistical Society*, 85(1), 87-94.

Cramer, H. (1946). *Mathematical Methods of Statistics*. Princeton.

test_correlation	<i>Test correlation between two numeric variables</i>
------------------	---

Description

Test correlation between two numeric variables

Usage

```
test_correlation(
  formula,
  data,
  y = NULL,
  method = c("auto", "pearson", "spearman", "kendall"),
  alpha = 0.05,
  plot = TRUE,
  na.rm = TRUE
)
```

Arguments

formula	A formula such as $y \sim x$, or a data frame when using pipe/data-first style.
data	A data frame, or a first numeric column when using data-first style.
y	Second numeric column. Optional when using formula style.
method	Correlation method or "auto".
alpha	Significance level.
plot	Logical; include a <code>ggplot</code> object.
na.rm	Logical; remove missing values.

Value

A testflow object with class `testflow_correlation`. The object is a list containing the cleaned complete-case data, numeric descriptives, assumption checks, recommended correlation method, primary correlation test with null hypothesis, Pearson/Spearman/Kendall results, a correlation table, effect size, optional ggplot, original call, and report text.

References

Pearson, K. (1895). Notes on regression and inheritance in the case of two parents. *Proceedings of the Royal Society of London*, 58, 240-242.

Spearman, C. (1904). The proof and measurement of association between two things. *The American Journal of Psychology*, 15(1), 72-101.

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika*, 30(1/2), 81-93.

`test_correlation_matrix`*Test a correlation matrix*

Description

Test a correlation matrix

Usage

```
test_correlation_matrix(  
  data,  
  vars,  
  method = c("spearman", "pearson", "kendall"),  
  alpha = 0.05,  
  plot = TRUE,  
  na.rm = TRUE  
)
```

Arguments

<code>data</code>	A data frame.
<code>vars</code>	Numeric columns.
<code>method</code>	Correlation method.
<code>alpha</code>	Significance level.
<code>plot</code>	Logical; include a ggplot object.
<code>na.rm</code>	Logical; remove missing values.

Value

A `testflow` object with class `testflow_correlation_matrix`. The object is a list containing the cleaned data, numeric descriptives, screening assumptions, selected correlation-matrix method, pairwise correlation matrix, pairwise p-value table, maximum absolute correlation as an effect-size summary, optional heatmap `ggplot`, original call, and report text.

<code>test_factorial</code>	<i>Run a factorial ANOVA workflow</i>
-----------------------------	---------------------------------------

Description

Run a factorial ANOVA workflow

Usage

```
test_factorial(
  formula,
  data,
  factors = NULL,
  alpha = 0.05,
  type = 2,
  plot = TRUE,
  na.rm = TRUE
)
```

Arguments

<code>formula</code>	A formula such as <code>outcome ~ factor1 * factor2</code> , or a data frame when using data-first style.
<code>data</code>	A data frame, or the outcome column when using data-first style.
<code>factors</code>	Factor columns selected with <code>tidyselect</code> syntax. Optional when using formula style.
<code>alpha</code>	Significance level.
<code>type</code>	ANOVA type placeholder for future car integration.
<code>plot</code>	Logical; include a <code>ggplot</code> object.
<code>na.rm</code>	Logical; remove missing values.

Value

A `testflow` object with class `testflow_factorial`. The object is a list containing the cleaned data, descriptive statistics, residual and variance assumption checks, recommended factorial ANOVA, primary ANOVA term result with null hypothesis, ANOVA table, effect size, optional `ggplot`, original call, and report text.

References

- Fisher, R. A. (1925). *Statistical Methods for Research Workers*. Oliver and Boyd.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum.

test_groups	<i>Compare a numeric outcome across more than two groups</i>
-------------	--

Description

Compare a numeric outcome across more than two groups

Usage

```
test_groups(
  formula,
  data,
  group = NULL,
  alpha = 0.05,
  posthoc = TRUE,
  plot = TRUE,
  na.rm = TRUE
)
```

Arguments

formula	A formula such as <code>outcome ~ group</code> , or a data frame when using pipe/data-first style.
data	A data frame, or an outcome column when using data-first style.
group	Grouping column. Optional when using formula style.
alpha	Significance level.
posthoc	Logical; compute post-hoc comparisons.
plot	Logical; include a ggplot object.
na.rm	Logical; remove missing values.

Value

A `testflow` object with class `testflow_groups`. The object is a list containing the cleaned data, grouped descriptive statistics, assumption checks, recommended omnibus test, primary test result with null hypothesis, alternative omnibus results, post-hoc comparisons when requested, effect size, optional ggplot, original call, and report text.

test_multinomial	<i>Test multinomial goodness of fit</i>
------------------	---

Description

Test multinomial goodness of fit

Usage

```
test_multinomial(  
  data,  
  outcome,  
  p = NULL,  
  alpha = 0.05,  
  plot = TRUE,  
  na.rm = TRUE  
)
```

Arguments

data	A data frame.
outcome	Categorical outcome column.
p	Expected probabilities, or NULL for equal probabilities.
alpha	Significance level.
plot	Logical; include a ggplot object.
na.rm	Logical; remove missing values.

Value

A testflow object with class `testflow_multinomial`. The object is a list containing the cleaned data, categorical descriptives, assumption checks, recommended goodness-of-fit test, primary chi-square result with null hypothesis, pairwise binomial checks, effect-size summary, optional ggplot, original call, and report text.

References

Pearson, K. (1900). On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *Philosophical Magazine*, 50(302), 157-175.

test_one_sample	<i>Test one numeric sample against a reference value</i>
-----------------	--

Description

Test one numeric sample against a reference value

Usage

```
test_one_sample(  
  data,  
  outcome,  
  mu = 0,  
  alternative = c("two.sided", "less", "greater"),  
  alpha = 0.05,  
  plot = TRUE,  
  na.rm = TRUE  
)
```

Arguments

data	A data frame.
outcome	Numeric outcome column.
mu	Reference value.
alternative	Alternative hypothesis.
alpha	Significance level.
plot	Logical; include a ggplot object.
na.rm	Logical; remove missing values.

Value

A `testflow` object with class `testflow_one_sample`. The object is a list containing the cleaned data, descriptive statistics, assumption checks, recommended test, primary test result with null hypothesis, alternative test results, effect size, optional `ggplot`, original call, and report text.

References

- Gosset, W. S. (1908). The probable error of a mean. *Biometrika*, 6(1), 1-25.
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6), 80-83.

test_outliers	<i>Detect numeric outliers</i>
---------------	--------------------------------

Description

Detect numeric outliers

Usage

```
test_outliers(
  formula,
  data,
  group = NULL,
  method = c("iqr", "mahalanobis", "both"),
  plot = TRUE,
  na.rm = TRUE
)
```

Arguments

formula	Numeric columns to screen for outliers.
data	A data frame.
group	Optional grouping column.
method	Outlier method.
plot	Logical; include a ggplot object.
na.rm	Logical; remove missing values.

Value

A testflow object with class testflow_outliers. The object is a list containing the cleaned data, numeric descriptives, screening assumptions, selected outlier-screening method, flagged IQR and/or Mahalanobis rows, outlier-count summary, optional ggplot, original call, and report text. This is a screening workflow, not a single hypothesis test.

test_paired	<i>Compare paired before and after numeric measurements</i>
-------------	---

Description

Compare paired before and after numeric measurements

Usage

```
test_paired(
  formula,
  data,
  after = NULL,
  alternative = c("two.sided", "less", "greater"),
  alpha = 0.05,
  plot = TRUE,
  na.rm = TRUE
)
```

Arguments

formula	A formula such as after ~ before, or a data frame when using data-first style.
data	A data frame, or the before column when using data-first style.
after	After column. Optional when using formula style.
alternative	Alternative hypothesis.
alpha	Significance level.
plot	Logical; include a ggplot object.
na.rm	Logical; remove missing values.

Value

A testflow object with class `testflow_paired`. The object is a list containing the cleaned paired data, paired difference, descriptive statistics, assumption checks, recommended paired test, primary test result with null hypothesis, alternative test results, effect size, optional ggplot, original call, and report text.

```
test_paired_categorical
```

Test paired categorical measurements

Description

Test paired categorical measurements

Usage

```
test_paired_categorical(
  data,
  before,
  after,
  alpha = 0.05,
  plot = TRUE,
  na.rm = TRUE
)
```

Arguments

data	A data frame.
before	Before categorical column.
after	After categorical column.
alpha	Significance level.
plot	Logical; include a ggplot object.
na.rm	Logical; remove missing values.

Value

A testflow object with class `testflow_paired_categorical`. The object is a list containing the cleaned paired categorical data, categorical descriptives, assumption checks, McNemar test result, discordant-pair table, optional ggplot, original call, and report text.

test_proportion	<i>Test a one-sample proportion</i>
-----------------	-------------------------------------

Description

Test a one-sample proportion

Usage

```
test_proportion(
  data,
  outcome,
  success,
  p = 0.5,
  alpha = 0.05,
  plot = TRUE,
  na.rm = TRUE
)
```

Arguments

data	A data frame.
outcome	Categorical outcome column.
success	Value counted as success.
p	Reference probability.
alpha	Significance level.
plot	Logical; include a ggplot object.
na.rm	Logical; remove missing values.

Value

A `testflow` object with class `testflow_proportion`. The object is a list containing the cleaned data, categorical descriptives, assumption checks, recommended exact or approximate one-sample proportion test, primary test result with null hypothesis, alternative exact and approximate results, observed proportion summary, optional `ggplot`, original call, and report text.

References

Clopper, C. J., & Pearson, E. S. (1934). The use of confidence or fiducial limits illustrated in the case of the binomial. *Biometrika*, 26(4), 404-413.

test_repeated	<i>Run a repeated-measures workflow from wide data</i>
---------------	--

Description

Run a repeated-measures workflow from wide data

Usage

```
test_repeated(
  data,
  measures,
  id = NULL,
  between = NULL,
  alpha = 0.05,
  plot = TRUE,
  na.rm = TRUE
)
```

Arguments

<code>data</code>	A data frame.
<code>measures</code>	Repeated numeric columns selected with <code>tidyselect</code> syntax.
<code>id</code>	Optional subject identifier.
<code>between</code>	Optional between-subject factor.
<code>alpha</code>	Significance level.
<code>plot</code>	Logical; include a <code>ggplot</code> object.
<code>na.rm</code>	Logical; remove missing values.

Value

A `testflow` object with class `testflow_repeated`. The object is a list containing long-form repeated-measures data, numeric descriptives, assumption checks, recommended repeated-measures ANOVA or Friedman test, primary test result with null hypothesis, alternative repeated-measures results, post-hoc paired comparisons, effect size, optional `ggplot`, original call, and report text.

References

- Fisher, R. A. (1925). *Statistical Methods for Research Workers*. Oliver and Boyd.
- Friedman, M. (1937). The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *Journal of the American Statistical Association*, 32(200), 675-701.
- Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6), 80-83.
- Girden, E. R. (1992). *ANOVA: Repeated Measures*. Sage.

test_repeated_categorical

Test repeated categorical measurements

Description

Test repeated categorical measurements

Usage

```
test_repeated_categorical(  
  data,  
  measures,  
  id = NULL,  
  alpha = 0.05,  
  plot = TRUE,  
  na.rm = TRUE  
)
```

Arguments

data	A data frame.
measures	Repeated binary columns selected with tidyselect syntax.
id	Optional subject identifier.
alpha	Significance level.
plot	Logical; include a ggplot object.
na.rm	Logical; remove missing values.

Value

A testflow object with class `testflow_repeated_categorical`. The object is a list containing the cleaned repeated categorical data, descriptive counts, assumption checks, Cochran Q test result with null hypothesis, pairwise McNemar post-hoc comparisons, effect size, optional ggplot, original call, and report text.

References

- Cochran, W. G. (1950). The comparison of percentages in matched samples. *Biometrika*, 37(3/4), 256-266.
- McNemar, Q. (1947). Note on the sampling error of the difference between correlated proportions or percentages. *Psychometrika*, 12(2), 153-157.

test_repeated_long	<i>Run a repeated-measures workflow from long data</i>
--------------------	--

Description

Run a repeated-measures workflow from long data

Usage

```
test_repeated_long(
  data,
  outcome,
  within,
  id,
  between = NULL,
  alpha = 0.05,
  plot = TRUE,
  na.rm = TRUE
)
```

Arguments

data	A data frame.
outcome	Numeric outcome column.
within	Within-subject time/condition column.
id	Subject identifier column.
between	Optional between-subject factor.
alpha	Significance level.
plot	Logical; include a ggplot object.
na.rm	Logical; remove missing values.

Value

A testflow object with class `testflow_repeated`. The object is a list containing the cleaned long-format data, numeric descriptives, assumption checks, recommended repeated-measures ANOVA or Friedman test, primary test result with null hypothesis, alternative repeated-measures results, post-hoc paired comparisons, effect size, optional ggplot, original call, and report text.

test_two_groups	<i>Compare a numeric outcome between two independent groups</i>
-----------------	---

Description

Compare a numeric outcome between two independent groups

Usage

```
test_two_groups(  
  formula,  
  data,  
  group = NULL,  
  alternative = c("two.sided", "less", "greater"),  
  alpha = 0.05,  
  plot = TRUE,  
  na.rm = TRUE  
)
```

Arguments

formula	A formula such as <code>outcome ~ group</code> , or a data frame when using pipe/data-first style.
data	A data frame, or an outcome column when using data-first style.
group	Two-level grouping column. Optional when using formula style.
alternative	Alternative hypothesis.
alpha	Significance level.
plot	Logical; include a ggplot object.
na.rm	Logical; remove missing values.

Value

A `testflow` object with class `testflow_two_groups`. The object is a list containing the cleaned data, descriptive statistics by group, assumption checks, recommended test, primary test result with null hypothesis, alternative test results, effect size, optional ggplot, original call, and report text.

Index

`as_tibble`, 2

`make_cardio_data`, 3

`plot.testflow`, 3

`print.testflow`, 4

`print.testflow()`, 5

`report`, 4

`report()`, 5

`report_test`, 5

`summary.testflow`, 5

`sumtab`, 6

`test_categorical`, 7

`test_correlation`, 8

`test_correlation_matrix`, 9

`test_factorial`, 10

`test_groups`, 11

`test_multinomial`, 12

`test_one_sample`, 13

`test_outliers`, 14

`test_paired`, 14

`test_paired_categorical`, 15

`test_proportion`, 16

`test_repeated`, 17

`test_repeated_categorical`, 18

`test_repeated_long`, 19

`test_two_groups`, 20