Package 'DataSum'

May 7, 2026

Title Comprehensive Data Summarization for Statistical Analysis

Version 0.1.1

Author Immad Ahmad Shah [aut],
Uzair Javid Khan [aut, cre],
Sukhdev Mishra [aut]

Maintainer Uzair Javid Khan <uzairkhan11w@gmail.com>

Description Summarizes data frames by calculating various statistics including central tendency, dispersion, shape, and normality diagnostics. Handles numeric, character, and factor columns with NA-aware computations.

License GPL-3

Encoding UTF-8

RoxygenNote 7.3.3

Imports moments, dplyr, nortest, stats

Suggests testthat (>= 3.0.0)

Config/testthat/edition 3

URL https://github.com/Uzairkhan11w/DataSum

BugReports https://github.com/Uzairkhan11w/DataSum/issues

NeedsCompilation no

Repository CRAN

Date/Publication 2025-11-12 18:00:21 UTC

Contents

DataSumm *Summarize an Entire Data Frame*

Description

Applies `Datum()` to each column of a data frame and binds the results.

Usage

```
DataSumm(data)
```

Arguments

`data`	A data frame (tibble is also ok).

Value

A data frame, one row per input column.

Examples

```
DataSumm(iris)
```

Datum *Summarize a Single Vector*

Description

Summarizes a single vector by calculating a consistent set of statistics. Numeric vectors include mean/median/variance/sd/min/max/range/skewness/kurtosis and a normality decision. Character/factor vectors report the mode only.

Usage

```
Datum(data)
```

Arguments

`data`	A numeric, character, factor, or other vector.

Value

A one-row `data.frame` with summary statistics.

Examples

```
Datum(rnorm(100))
Datum(factor(sample(letters[1:3], 20, TRUE)))
```

getmode *Get Mode of a Vector (numeric/character/factor)*

Description

Returns the mode(s) of a vector. By default, returns a single string with all modes collapsed by comma when there are ties. If you need only one mode, set `collapse = FALSE` to return the first mode deterministically.

Usage

```
getmode(x, collapse = TRUE)
```

Arguments

`x`	A vector (numeric, character, factor, etc.).
`collapse`	Logical; if TRUE (default), return all modes as a single comma-separated string. If FALSE, return the first mode only.

Value

A single value (first mode) or a comma-separated string of modes.

Examples

```
getmode(c(1, 2, 2, 3, 4))
getmode(c("a", "b", "b", "a"), collapse = TRUE)
getmode(c("a", "b", "b", "a"), collapse = FALSE)
```

shapiro_normality_test *Shapiro/Anderson-Darling Normality Decision*

Description

Performs Shapiro-Wilk for sample sizes between 3 and 5000 (inclusive), otherwise uses Anderson–Darling. Returns "Normal" if $p > 0.05$, else "Not Normal".

Usage

```
shapiro_normality_test(data)
```

Arguments

`data`	A numeric vector.

Value

Character scalar: "Normal", "Not Normal", or `NA` if not applicable.

Examples

```
shapiro_normality_test(rnorm(100))
```

Index