Package ‘wrangle’

April 25, 2019

Type Package
Title A Systematic Data Wrangling Idiom
Version 0.5.2
Author Tim Bergsma
Maintainer Tim Bergsma <bergsmat@gmail.com>
Description Supports systematic scrutiny, modification, and integration of data. The function status() counts rows that have missing values in grouping columns (returned by na()), have non-unique combinations of grouping columns (returned by dup()), and that are not locally sorted (returned by unsorted()). Functions enumerate() and itemize() give sorted unique combinations of columns, with or without occurrence counts, respectively. Function ignore() drops columns in x that are present in y, and informative() drops columns in x that are entirely NA; constant() returns values that are constant, given a key. Data that have defined unique combinations of grouping values behave more predictably during merge operations.
License GPL-3
LazyData TRUE
Imports dplyr, tidyr, lazyeval, magrittr, rlang
RoxygenNote 6.0.1
NeedsCompilation no
Repository CRAN
Date/Publication 2019-04-25 05:00:03 UTC

R topics documented:

constant ............................................................... 2
costant.data.frame .................................................. 3
costant.grouped_df .................................................. 3
detect ................................................................. 4
dup ................................................................. 5
dup.grouped_df ..................................................... 5
dupGroups .......................................................... 6
Identify Constant Features of an Object

Description

Identifies constant features of an object. Generic, with methods for data.frame and grouped_df.

Usage

constant(x, ...)

Arguments

x object

... passed arguments

See Also

Other constant: constant.data.frame, constant.grouped_df
constant.data.frame  Identify Constant Features of a Data Frame

Description

Returns columns of a data.frame whose values do not vary within subsets defined by columns named in ....

Usage

## S3 method for class 'data.frame'
constant(x, ...)

Arguments

x  object

... grouping columns

Value

data.frame

See Also

Other constant: constant.grouped_df, constant

Examples

constant(Theoph)
class(Theoph, Subject) # Wt Dose
Theoph$Study <- 1
class(Theoph)
class(Theoph, Study)
class(Theoph, Study, Subject)

constant.grouped_df  Identify Constant Features of a Grouped Data Frame

Description

Returns columns of a grouped_df whose values do not vary within subsets defined by groups. If any grouping arguments (dots) are supplied, existing groups are over-ridden.

Usage

## S3 method for class 'grouped_df'
constant(x, ...)


Arguments

x          object
...        grouping columns

Value

grouped data.frame

See Also

Other constant: \texttt{constant.data.frame, constant}

---

detect \> Sort column subsets.

Description

Sort column subsets.

Usage

detect(x, ...)

Arguments

x          data.frame
...        columns to sort

Value

grouped\_df

See Also

Other util: \texttt{enumerate, itemize, static}
**dup**

*Show duplicate or duplicated elements.*

**Description**

Shows duplicate or duplicated elements.

**Usage**

```r
dup(x, ...)```

**Arguments**

- `x` object of dispatch
- `...` other arguments

**See Also**

`dup.grouped_df` `na` `weak` `unsorted`

Other `dup`: `dup.grouped_df`

---

**dup.grouped_df**

*Show records with duplicate or duplicated values of grouping variables.*

**Description**

Shows records with duplicate or duplicated values of grouping variables.

**Usage**

```r
## S3 method for class 'grouped_df'
dup(x, ...)
```

**Arguments**

- `x` data.frame
- `...` ignored

**Value**

`grouped_df`

**See Also**

Other `dup`: `dup`
**Examples**

```r
library(dplyr)
dup(group_by(mtcars, mpg))
```

---

**Description**

Calculates dupGroups.

**Usage**

```r
dupGroups(x, ...)
```

**Arguments**

- `x`: object of dispatch
- `...`: other arguments

**See Also**

Other dupGroups: `dupGroups.grouped_df`

---

**Description**

Counts records with duplicate or duplicated values of grouping variables. If b follows a and is the same, then b is a duplicate, a is duplicated, and both are shown.

**Usage**

```r
## S3 method for class 'grouped_df'
dupGroups(x, ...)
```

**Arguments**

- `x`: data.frame
- `...`: ignored

**Value**

`grouped_df`
enumerate

See Also

Other dupGroups: dupGroups

---

**enumerate**

Count unique combinations of items in specified columns.

**Description**

Counts unique combinations of items in specified columns (unquoted).

**Usage**

enumerate(x, ...)

**Arguments**

- **x**
  - data.frame
- **...**
  - columns to show

**Value**

grouped_df

**See Also**

Other util: detect, itemize, static

**Examples**

enumerate(mtcars, cyl, gear, carb)

---

group_by_all

Group by all columns.

**Description**

Groups by all columns.

**Usage**

group_by_all(x, ...)

**Arguments**

- **x**
  - data.frame
- **...**
  - ignored
Value
grouped_df

ignore

Drop columns in x that are present in y.

Description
Drops columns in x that are present in y.

Usage
ignore(x, y, ...)

Arguments

x  data.frame
y  data.frame
...  ignored

Value
data.frame

informative

Drop columns in x that are entirely NA.

Description
Drops columns in x that are entirely NA.

Usage
informative(x, ...)

Arguments

x  object of dispatch
...  passed

See Also
informative.data.frame
Other informative: informative.data.frame
Examples

```r
head(theoph)
Theoph$Dose <- NA
head(informative(Theoph))
```

---

**informative.data.frame**

*Drop columns in x that are entirely NA.*

---

**Description**

Drops columns in x that are entirely NA.

**Usage**

```r
## S3 method for class 'data.frame'
informative(x, ...)
```

**Arguments**

- **x** : data.frame
- **...** : ignored

**Value**

data.frame

**See Also**

Other informative: `informative`

---

**itemize**

*Show unique combinations of items in specified columns*

---

**Description**

Shows unique combinations of items in specified columns (unquoted).

**Usage**

```r
itemize(x, ...)
```

**Arguments**

- **x** : data.frame
- **...** : columns to show
Value

grouped_df

See Also

Other util: detect, enumerate, static

Examples

itemize(mtcars, cyl, gear, carb)

<table>
<thead>
<tr>
<th>key</th>
<th>Fetch the key.</th>
</tr>
</thead>
</table>

Description

Fetches the key of an object.

Usage

key(x, ...)

Arguments

x          object of dispatch
...
other arguments

See Also

Other key: key.grouped_df

<table>
<thead>
<tr>
<th>key.grouped_df</th>
<th>Fetch the key for a grouped_df as character vector</th>
</tr>
</thead>
</table>

Description

Fetches the key for a grouped_df as character vector.

Usage

## S3 method for class 'grouped_df'
key(x, ...)

key.grouped_df
Arguments

- `x`: data.frame
- `...`: columns to show

Value

- character

See Also

Other key: `key`

---

### na

**Show na elements.**

Description

Shows na elements.

Usage

`na(x, ...)`

Arguments

- `x`: object of dispatch
- `...`: other arguments

See Also

- `na.grouped_df`
- `dup` weak unsorted

Other na: `na.grouped_df`
Description

Shows records with NA values of grouping variables.

Usage

```r
## S3 method for class 'grouped_df'
na(x, ...)
```

Arguments

- `x`: data.frame
- `...`: ignored

Value

grouped_df

See Also

Other na: `na`

Description

Calculates naGroups.

Usage

```r
naGroups(x, ...)
```

Arguments

- `x`: object of dispatch
- `...`: other arguments

See Also

Other naGroups: `naGroups.grouped_df`
naGroups.grouped_df  Count records with NA values of grouping variables.

Description
Counts records with NA values of grouping variables.

Usage
## S3 method for class 'grouped_df'
naGroups(x, ...)

Arguments

- x  data.frame
- ...  ignored

Value
numeric

See Also
Other naGroups: naGroups

sort.grouped_df  Arrange by groups.

Description
As of 0.5, dplyr::arrange ignores groups. This function gives the old behavior as a method for

Usage
## S3 method for class 'grouped_df'
sort(x, decreasing = FALSE, ...)

Arguments

- x  grouped_df
- decreasing  logical (ignored)
- ...  further sort criteria
**Value**

```
grouped_df
```

**Examples**

```
library(dplyr)
head(sort(group_by(Theoph, Subject, Time)))
```

---

**static**  
*Find unique records for subset of columns with one unique value.*

**Description**

Finds unique records for subset of columns with one unique value.

**Usage**

```
static(x, ...)
```

**Arguments**

- `x`  
  *data.frame*
- `...`  
  *ignored*

**Value**

```
data.frame
```

**See Also**

Other util: `detect, enumerate, itemize`

---

**status**  
*Report status.*

**Description**

Reports the status of an object.

**Usage**

```
status(x, ...)
```
status.grouped_df

Arguments

x  object of dispatch
...

other arguments

See Also

Other status: status.grouped_df

Examples

library(dplyr)
status(group_by(Theoph, Subject, Time))

status.grouped_df  Report status with respect to grouping variables.

Description

Reports status with respect to grouping variables.

Usage

## S3 method for class 'grouped_df'
status(x, ...)

Arguments

x  data.frame
...

ignored

Value

returns x invisibly

See Also

na dup unsorted informative ignore itemize enumerate sort.grouped_df

Other status: status

Examples

library(dplyr)
status(group_by(Theoph, Subject, Time))
unsorted

Show unsorted elements.

Description

Shows unsorted elements.

Usage

unsorted(x, ...)

Arguments

x object of dispatch
... other arguments

See Also

unsorted.grouped_df

Other unsorted: unsorted.grouped_df

unsorted.grouped_df

Find records whose relative positions would change if sorted.

Description

Finds records whose relative positions would change if sorted, i.e. records that would not have the same nearest neighbors (before and after).

Usage

## S3 method for class 'grouped_df'
unsorted(x, ...)

Arguments

x data.frame
... ignored

Value

grouped_df

See Also

na dup

Other unsorted: unsorted
weak

Show na, duplicate, or duplicated elements.

Description

Shows na, duplicate, or duplicated elements.

Usage

weak(x, ...)

Arguments

x  object of dispatch
...
other arguments

See Also

weak.grouped_df

Other weak: weak.grouped_df

weak.grouped_df

Show records with NA, duplicate or duplicated values of grouping variables.

Description

Shows records with NA, duplicate or duplicated values of grouping variables.

Usage

## S3 method for class 'grouped_df'
weak(x, ...)

Arguments

x  data.frame
...
ignore

Value

grouped_df

See Also

Other weak: weak
Index

constant, 2, 3, 4
constant.data.frame, 2, 3, 4
constant.grouped_df, 2, 3, 3
detect, 4, 7, 10, 14
dup, 5, 5, 11, 15, 16
dup.grouped_df, 5, 5
dupGroups, 6, 7
dupGroups.grouped_df, 6, 6
enumerate, 4, 7, 10, 14, 15
group_by_all, 7
ignore, 8, 15
informative, 8, 9, 15
informative.data.frame, 8, 9
itemize, 4, 7, 9, 14, 15
key, 10, 11
key.grouped_df, 10, 10
na, 5, 11, 12, 15, 16
na.grouped_df, 11, 12
naGroups, 12, 13
naGroups.grouped_df, 12, 13
sort.grouped_df, 13, 15
static, 4, 7, 10, 14
status, 14, 15
status.grouped_df, 15, 15
unsorted, 5, 11, 15, 16, 16
unsorted.grouped_df, 16, 16
weak, 5, 11, 17, 17
weak.grouped_df, 17, 17
wrangle (status.grouped_df), 15