Package 'makedummies'

May 8, 2026

Type Package

Title Create Dummy Variables from Categorical Data

Version 1.2.1

Date 2019-08-04

Description Create dummy variables from categorical data.
This package can convert categorical data (factor and ordered) into dummy variables and handle multiple columns simultaneously.
This package enables to select whether a dummy variable for base group is included (for principal component analysis/factor analysis) or excluded (for regression analysis) by an option.
'makedummies' function accepts 'data.frame', 'matrix', and 'tbl' (tibble) class (by 'tibble' package).
'matrix' class data is automatically converted to 'data.frame' class.

Imports tibble

License GPL-2

URL https://github.com/toshi-ara/makedummies

BugReports https://github.com/toshi-ara/makedummies/issues/

RoxygenNote 6.1.1

NeedsCompilation no

Author Toshiaki Ara [aut, cre]

Maintainer Toshiaki Ara <toshiaki.ara@gmail.com>

Repository CRAN

Date/Publication 2019-08-04 11:00:02 UTC

Contents

makedummies *Create Dummy Variables from Categorical Data*

Description

Create dummy variables from categorical data. This package can convert categorical data (factor and ordered) into dummy variables and handle multiple columns simultaneously. This package enables to select whether a dummy variable for base group is included (for principal component analysis/factor analysis) or excluded (for regression analysis) by an option. `makedummies` function accepts `data.frame`, `matrix`, and `tbl` (tibble) class (by `tibble` package). `matrix` class data is automatically converted to `data.frame` class.

Usage

```
makedummies(dat, ...)

## Default S3 method:
makedummies(dat, basal_level = FALSE, col = NULL,
  numerical = NULL, as.is = NULL, ...)

## S3 method for class 'matrix'
makedummies(dat, ...)

## S3 method for class 'tbl'
makedummies(dat, basal_level = FALSE, col = NULL,
  numerical = NULL, as.is = NULL, ...)
```

Arguments

`dat`	data of `data.frame`, `matrix`, or `tbl` class
`...`	arguments to makedummies.data.frame (`tbl` class)
`basal_level`	logical **TRUE** : include a dummy variable for base group **FALSE** (default) : exclude a dummy variable for base group
`col`	Columns vector (all columns are used if `NULL` is given)
`numerical`	Columns vector converting from `factor/ordered` to `numeric` (ignore if column is `numeric`)
`as.is`	Columns vector not converting

Value

return as `data.frame` or `tbl` class

Note

Pull Request #1 (add column name when when columns has binary value) (https://github.com/toshi-ara/makedummies/pull/1). Thanks to Kohki YAMAGIWA for the contribution.

Examples

```
#### 'data.frame' class
## factor
dat <- data.frame(x = factor(rep(c("a", "b", "c"), each = 3)))
dat$x
makedummies(dat)

## ordered
dat <- data.frame(x = factor(rep(c("a", "b", "c"), each = 3)))
dat$x <- ordered(dat$x, levels = c("a" ,"c" ,"b"))
dat$x
makedummies(dat)

## numeric
dat <- data.frame(x = rep(1:3, each = 3))
makedummies(dat)

## factor and numeric
dat <- data.frame(
    x = factor(rep(c("a", "b", "c"), each = 3)),
    y = rep(1:3, each = 3)
)
makedummies(dat)

## factors
dat <- data.frame(
    x = factor(rep(c("a", "b", "c"), each = 3)),
    y = factor(rep(1:3, each = 3))
)
makedummies(dat)

## data including NA

dat <- data.frame(
    x = factor(rep(c("a", "b", "c"), each = 3)),
    y = rep(1:3, each = 3)
)
dat$x[4] <- NA; dat$y[6] <- NA
makedummies(dat)

## "col" option
dat <- data.frame(
    x = factor(rep(c("a", "b", "c"), each = 3)),
    y = factor(rep(1:3, each = 3))
)
makedummies(dat, col = "x")

## "numerical" option
dat <- data.frame(
    x = factor(rep(c("a", "b", "c"), each = 3)),
    y = factor(rep(1:3, each = 3))
)
```

```
makedummies(dat, numeric = "x")

dat <- data.frame(
    x = factor(rep(c("a", "b", "c"), each = 3)),
    y = rep(4:6, each = 3)
)
dat$x <- ordered(dat$x, levels = c("a" ,"c" ,"b"))
dat
dat$x
makedummies(dat, numeric = c("x", "y"))

## "as.is" option
dat <- data.frame(
    x = factor(rep(c("a", "b", "c"), each = 3)),
    y = factor(rep(1:3, each = 3))
)
dat
makedummies(dat, as.is = "x")
makedummies(dat, as.is = c("x", "y"))

#### 'tibble' class
if (require(tibble)) {
  dat <- as_tibble(iris)
  makedummies(dat[46:55,], col = "Species", basal_level = TRUE)

  # non-standard variable name
  dat2 <- tibble(
      `1` = factor(rep(c("c", "a", "b"), each = 3)),
      `@` = factor(rep(1:3, each = 3)),
      `&` = rep(4:6, each = 3)
  )
  dat2

  makedummies(dat2, basal_level = TRUE)
  makedummies(dat2, as.is = "@", basal_level = TRUE)
  makedummies(dat2, numerical = "1", basal_level = TRUE)
}
```

Index