Package 'AggregateR'

November 23, 2015

Type Package

Title Aggregate Numeric, Date and Categorical Variables by an ID

Version 0.0.2

Date 2015-11-20

Imports dummy, stats, utils, methods

Author Matthias Bogaert, Michel Ballings, Dirk Van den Poel

Maintainer Matthias Bogaert <matthias.bogaert@ugent.be>

Description Convenience functions for aggregating data frame. Currently mean, sum and variance are supported. For Date variables, recency and duration are supported. There is also support for dummy variables in predictive contexts.

License GPL (>= 2)

LazyData TRUE

NeedsCompilation no

Repository CRAN

Date/Publication 2015-11-23 12:33:51

R topics documented:

Aggregate	*Aggregate numeric,Date and categorical variables by an ID*

Description

The Aggregate function (not to be confounded with aggregate) prepares a data frame for merging by computing the sum, mean and variance of all continuous (integer and numeric) variables by a given ID variable. For all categorical variabes (character and factor), it creates dummies and subsequently computes the sum by a given ID variable. For all Date variables, it computes recency and duration by a given ID and end date variable. This functions aims at maximum information extraction with a minimum amount of code.

Usage

```
Aggregate(x, by, end_ind = Sys.Date(), format = "%Y-%m-%d", ...)
```

Arguments

x	A data frame without the ID. Categorical variables have to be of type character or factor and continuous variables have to be of type integer or numeric.
by	A vector containing IDs.
end_ind	A Date object, or something which can be coerced by as.Date(origin, ...) to such an object. If not specified, we take the Sys.Date() as end date.
format	A character string. If not specified, the ISO 8601 international standard which expresses a day "%Y-%m-%d" is taken.
...	Extra parameters to be passed to the dummy function in the dummy package.

Value

A data frame with the aforementioned variables aggregated by the given ID variables

Author(s)

Authors: Matthias Bogaert, Michel Ballings, Dirk Van den Poel, Maintainer: <matthias.bogaert@UGent.be>

Examples

```
# Example
# Create some data
data <- data.frame(V1=as.factor(c('yes','no','no','yes','yes','yes','yes')),
                   V2=as.character(c(1,2,3,4,4,4,4)),V3=c(1:7),V4=as.numeric(c(7:1)),
               V5 = as.Date(as.Date('2014-12-03'):as.Date('2014-12-09'), origin = "1970-01-01")
                   )
ID=as.character(c(1,1,1,1,2,2,2))
Aggregate(x=data,by=ID)

# Examples of how to use the ... argument. See package dummy for details.
# library(dummmy)
# Aggregate(x=data,by=ID,object=categories(data))
# Aggregate(x=data,by=ID,p=2)
```

Index