
`chileancredit` *Chilean Credit Data*

Description

A simulated dataset where the target variable is fgood, which represents the binary status of default (0) and not default (1).

Format

Data frame with 10,000 rows and 22 columns.

Details

- fgood: Default (0), Not Default (1).
- cbs1: Credit score 1.
- cbs2: Credit score 2.
- cbs3: Credit score 3.
- cbinq: Number of inquiries.
- cbline: Number of credit lines.
- cbterm: Number of term loans.
- cblineut: Line utilization (0-100).
- cbtob: Number of years on file.
- cbdpd: Indicator of days past due on bureau (Yes, No).
- cbnew: Number of new loans.
- pmt: Type of payment (M: Manual, A: Autopay, P: Payroll).
- tob: Time on books (Years).
- dpd: Level of delinquency (No, Low, High).
- dep: Amount of deposits own by customer.
- dc: Number of debit card transactions.
- od: Number of overdrafts.

- home: Home ownership indicator (Yes, No).
- inc: Level of income.
- dd: Number of direct deposits per month.
- online: Indicator of active online (Yes, No).
- rnd: Random number to select testing and training samples.

`smbinning` *Optimal Binning for Scoring Modeling*

Description

Optimal Binning categorizes a numeric characteristic into bins for ulterior usage in scoring modeling. This process, also known as *supervised discretization*, utilizes Recursive Partitioning to categorize the numeric characteristic.
The especific algorithm is Conditional Inference Trees which initially excludes missing values (`NA`) to compute the cutpoints, adding them back later in the process for the calculation of the *Information Value*.

Usage

```
smbinning(df, y, x, p = 0.05)
```

Arguments

`df`	A data frame.
`y`	Binary response variable (0,1). Integer (`int`) is required. Name of `y` must not have a dot. Name "default" is not allowed.
`x`	Continuous characteristic. At least 5 different values. Value `Inf` is not allowed. Name of `x` must not have a dot.
`p`	Percentage of records per bin. Default 5% (0.05). This parameter only accepts values greater that 0.00 (0%) and lower than 0.50 (50%).

Value

The command `smbinning` generates and object containing the necessary info and utilities for binning. The user should save the output result so it can be used with `smbinning.plot`, `smbinning.sql`, and `smbinning.gen`.

Examples

```
# Load library and its dataset
library(smbinning) # Load package and its data

# Example: Optimal binning
result=smbinning(df=chileancredit,y="fgood",x="cbs1") # Run and save result
result$ivtable # Tabulation and Information Value
```