

Titanic help

Description

Titanic data with incomplete cases, passenger names, and other details removed.

Format

A data frame with 1046 observations on 6 variables.

Variable	Description
pclass	passenger class, unordered factor: 1st 2nd 3rd
survived	integer: 0 or 1
sex	unordered factor: male female
age	age in years, min 0.167 max 80.0
sibsp	number of siblings or spouses aboard, integer: 0 . . 8
parch	number of parents or children aboard, integer: 0 . . 6

This dataset is included in the earth package because it is a convenient vehicle for illustrating earth's GLM and factor handling.

The dataset was compiled by Frank Harrell and Robert Dawson: <http://biostat.mc.vanderbilt.edu/twiki/pub/Main/DataSets/titanic.html>

See also: <http://biostat.mc.vanderbilt.edu/twiki/pub/Main/DataSets/titanic3info.txt>.

For this version of the Titanic data, passenger details and incomplete cases were deleted and the name changed to etitanic to minimize confusion with other versions.

Note that survived is an integer (it should arguably be a logical).

References

Further details and analyses of the Titanic data may be found in: F. Harrell (2001) Regression Modeling Strategies with Applications to Linear Models, Logistic Regression, and Survival Analysis <http://biostat.mc.vanderbilt.edu/twiki/bin/view/Main/RmS>