Dataset Description (Stat2017)

Fu Xing

2017-Aug-08

1. Motor Trend Car Road Tests

Description

The data was extracted from the 1974 Motor Trend US magazine, and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973–74 models).

Usage

mtcars

Format

A data frame with 32 observations on 11 variables.

[, 1] mpg Miles/(US) gallon

[, 2] cyl Number of cylinders

[, 3] disp Displacement (cu.in.)

[, 4] hp Gross horsepower

[, 5] drat Rear axle ratio

[, 6] wt Weight (1000 lbs)

[, 7] qsec 1/4 mile time

[, 8] vs V/S

[, 9] am Transmission (0 = automatic, 1 = manual)

[,10] gear Number of forward gears

[,11] carb Number of carburetors

2. Arthritis Treatment Data

Description

Data from Koch & Edwards (1988) from a double-blind clinical trial investigating a new treatment for rheumatoid arthritis.

Usage

vcd::Arthritis

Format

A data frame with 84 observations and 5 variables.

ID: patient ID.

Treatment: factor indicating treatment (Placebo, Treated).

Sex: factor indicating sex (Female, Male).

Age: age of patient.

Improved: ordered factor indicating treatment outcome (None, Some, Marked).

3. US State Facts and Figures

Description

Data sets related to the 50 states of the United States of America.

Usage

state.abb

state.area

state.center

state.division

state.name

state.region

state.x77

Details

R currently contains the following "state" data sets. Note that all data are arranged according to alphabetical order of the state names.

state.abb: character vector of 2-letter abbreviations for the state names.

state.area: numeric vector of state areas (in square miles).

state.center: list with components named x and y giving the approximate geographic center of each state in negative longitude and latitude. Alaska and Hawaii are placed just off the West Coast.

state.division: factor giving state divisions (New England, Middle Atlantic, South Atlantic, East South Central, West South Central, East North Central, West North Central, Mountain, and Pacific).

state.name: character vector giving the full state names.

state.region: factor giving the region (Northeast, South, North Central, West) that each state belongs to.

state.x77: matrix with 50 rows and 8 columns giving the following statistics in the respective columns.

- Population: population estimate as of July 1, 1975
- Income: per capita income (1974)
- Illiteracy: illiteracy (1970, percent of population)
- Life Exp: life expectancy in years (1969–71)
- Murder: murder and non-negligent manslaughter rate per 100,000 population (1976)
- HS Grad: percent high-school graduates (1970)
- Frost: mean number of days with minimum temperature below freezing (1931–1960) in capital or large city
- Area: land area in square miles

4. The Effect of Punishment Regimes on Crime Rates

Description

Criminologists are interested in the effect of punishment regimes on crime rates. This has been studied using aggregate data on 47 states of the USA for 1960 given in this data frame. The variables seem to have been re-scaled to convenient numbers.

Usage

MASS::UScrime

Format

This data frame contains the following columns:

M: percentage of males aged 14–24.

So: indicator variable for a Southern state.

Ed: mean years of schooling.

Po1: police expenditure in 1960.

Po2: police expenditure in 1959.

LF: labour force participation rate.

M.F: number of males per 1000 females.

Pop: state population.

NW: number of non-whites per 1000 people.

U1: unemployment rate of urban males 14–24.

U2: unemployment rate of urban males 35–39.

GDP: gross domestic product per head.

Ineq: income inequality.

Prob: probability of imprisonment.

Time: average time served in state prisons.

y: rate of crimes in a particular category per head of population.

5. Average Heights and Weights for American Women

Description

This data set gives the average heights and weights for American women aged 30–39.

Usage

women

Format

A data frame with 15 observations on 2 variables.

[,1] height numeric Height (in)

[,2] weight numeric Weight (lbs)

6. Cholesterol Reduction Data Set

Description

Cholesterol reduction for five treatments.

Usage

multcomp::cholesterol

Format

This data frame contains the following variables

trt: treatment groups, a factor at levels 1time, 2times, 4times, drugD and drugE.

response: cholesterol reduction.

Details

A clinical study was conducted to assess the effect of three formulations of the same drug on reducing cholesterol. The formulations were 20mg at once (1time), 10mg twice a day (2times), and 5mg four times a day (4times). In addition, two competing drugs were used as control group (drugD and drugE). The purpose of the study was to find which of the formulations, if any, is efficacious and how these formulations compare with the existing drugs.

7. Litter Weights Data Set

Description

Dose response of litter weights in rats.

Usage

multcomp::litter

Format

This data frame contains the following variables

dose: dosages at four levels: 0, 5, 50, 500.

gesttime: gestation time as covariate.

number: number of animals in litter as covariate.

weight: response variable: average post-birth weights in the entire litter.

Details

Pregnant mice were divided into four groups and the compound in four different doses was administered during pregnancy. Their litters were evaluated for birth weights.

8. The Effect of Vitamin C on Tooth Growth in Guinea Pigs

Description

The response is the length of odontoblasts (cells responsible for tooth growth) in 60 guinea pigs. Each animal received one of three dose levels of vitamin C (0.5, 1, and 2 mg/day) by one of two delivery methods, (orange juice or ascorbic acid (a form of vitamin C and coded as VC).

Usage

ToothGrowth

Format

A data frame with 60 observations on 3 variables.

[,1] len numeric Tooth length

[,2] supp factor Supplement type (VC or OJ).

[,3] dose numeric Dose in milligrams/day

9. Carbon Dioxide Uptake in Grass Plants

Description

The CO2 data frame has 84 rows and 5 columns of data from an experiment on the cold tolerance of the grass species Echinochloa crus-galli.

Usage

CO2

Format

An object of class c("nfnGroupedData", "nfGroupedData", "groupedData", "data.frame") containing the following columns:

Plant: an ordered factor with levels Qn1 < Qn2 < Qn3 < ... < Mc1 giving a unique identifier for each plant.

Type: a factor with levels Quebec Mississippi giving the origin of the plant

Treatment: a factor with levels nonchilled chilled

conc: a numeric vector of ambient carbon dioxide concentrations (mL/L).

uptake: a numeric vector of carbon dioxide uptake rates (umol/m^2 sec).

Details

The CO2 uptake of six plants from Quebec and six plants from Mississippi was measured at several levels of ambient CO2 concentration. Half the plants of each type were chilled overnight before the experiment was conducted.

This dataset was originally part of package nlme, and that has methods (including for [, as.data.frame, plot and print) for its grouped-data classes.