

Package ‘smss’

July 23, 2025

Version 1.0-2

Title Datasets for Agresti and Finlay's ``Statistical Methods for the Social Sciences''

Suggests devtools

Description Datasets used in ``Statistical Methods for the Social Sciences'' (SMSS) by Alan Agresti and Barbara Finlay.

License GPL-3

NeedsCompilation no

Author Jeffrey B. Arnold [aut, cre],
Alan Agresti [cph],
Barbara Finlay [cph]

Maintainer Jeffrey B. Arnold <jeffrey.arnold@gmail.com>

Repository CRAN

Date/Publication 2015-10-31 00:30:27

Contents

smss-package	2
anorexia	2
birth.rates	3
crime2005	3
fertility.gdp	4
fl.crime	5
house.selling.price	5
house.selling.price.2	6
inc.ed.race.13p1	6
income.credit	7
mental.impairment	8
oecd.data	8
statewide.crime.2	9
student.survey	10
UNdata	11
us.pop.size	12
zagat	12

Index**14**

smss-package

*Data from Agresti "Statistical Methods for the Social Sciences"***Description**

Data-only package with the datasets used by the textbook, *Statistical Methods for the Social Sciences* by Alan Agresti and Barbara Finlay.

- Agresti, Alan and Barbara Finlay (2009), *Statistical Methods for the Social Sciences*, 4th ed., Pearson
- Agresti, Alan and Barbara Finlay (1997), *Statistical Methods for the Social Sciences*, 3rd ed., Pearson

The datasets included in this package are those available at <http://www.stat.ufl.edu/~aa/social/data.html>.

anorexia

*Anorexia Study data***Description**

Weights of Anorexic girls, before and after receiving one of three possible therapies. Thanks to Prof. Brian Everitt, Institute of Psychiatry, London, for supplying these data.

Format

data.frame with 72 observations of the following 4 variables,

subj Subject ID

therapy Therapy type. b = cognitive behavioural, f = family therapy, or c = control.

before Weight before treatment

after Weight after treatment

Source

<http://www.stat.ufl.edu/~aa/social/data.html>

Examples

```
data(anorexia)
summary(anorexia)
```

`birth.rates`*Birth Rate Data*

Description

Birth rates of several nations. Table 9.13 of 3rd edition.

Format

`data.frame` with 29 observations of the following 10 variables,

B crude birth rate (number of births per 1000 population size)

W women's economic activity (female labor force as percent of male)

C percent women using contraception

LI female adult literacy rate

LE female life expectancy

HDI human development index (which has components referring to life expectancy at birth, educational attainment, and income per capita)

GNP gross national product (per capita, in thousands of dollars)

N daily newspaper circulation per 100 people

T number of televisions per 100 people.

Source

<http://www.stat.ufl.edu/~aa/social/data.html>

Examples

```
data(birth.rates)
summary(birth.rates)
```

`crime2005`*2005 Statewise Crime*

Description

Data from the Statistical Abstract of the United States

Format

data.frame with 51 observations of the following 8 variables,

STATE U.S. state

VI violent crime rate (number of violent crimes per 100,000 population)

VI2 violent crime rate (number of violent crimes per 10,000 population)

MU murder rate

ME percent in metropolitan areas

WH percent white

HS percent high school graduates

PO percent below the poverty level.

Source

<http://www.stat.ufl.edu/~aa/social/data.html>

Examples

```
data(crime2005)
summary(crime2005)
```

fertility.gdp

Fertility and GDP Data

Description

Table 14.6 from the 4th edition.

Format

A data.frame with 38 observations on the following 3 variables,

Nation nation

Fertility Fertility rate

GDP GDP per capita

Source

<http://www.stat.ufl.edu/~aa/social/data.html>

Examples

```
data(fertility.gdp)
summary(fertility.gdp)
```

fl.crime*Florida Crime Data*

Description

Crime data for florida counties. Table 9.16 or 9.17 from the 4th edition. Source: Dr. Larry Winner, University of Florida.

Format

data.frame with 67 observations of the following 5 variables,

County county name

C crime rate

I median income

HS percent completing high school

U percent urban

Source

<http://www.stat.ufl.edu/~aa/social/data.html>

Examples

```
data(fl.crime)
summary(fl.crime)
```

house.selling.price*House Selling Price Data*

Description

Selling price of homes in Gainesville, Florida, fall 2006, from Alachua County public records. Excerpt in Table 9.4 of 4th edition.

Format

data.frame with 100 observations of the following 7 variables,

case observation id

Baths number of bathrooms

New whether new (1 = yes, 0 = no)

Price selling price (dollars)

Size size of home (square feet)

Source

<http://www.stat.ufl.edu/~aa/social/data.html>

Examples

```
data(house.selling.price)
summary(house.selling.price)
```

house.selling.price.2 *House Selling Price 2 Data*

Description

Selling price of homes in Gainesville, Florida, January 1996. Table 9.4 in 3rd edition. Data provided by Jane Myers, Coldwell-Banker Realty.

Format

data.frame with 93 observations of the following 5 variables,

P selling price

Ba number of bathrooms

New whether new (1 = yes, 0 = no)

Source

<http://www.stat.ufl.edu/~aa/social/data.html>

Examples

```
data(house.selling.price.2)
summary(house.selling.price.2)
```

inc.ed.race.13p1 *Income, Education and Race Data*

Description

Data on annual income, number of years of education, and racial-ethnic group (Black, Hispanic, White). Table 13.1 of 4th edition.

Format

data.frame with 80 observations of the following 5 variables,

- inc income (thousands of dollars)
- educ number of years of education (where 12 = high school graduate, 16 = college graduate)
- race racial-ethnic group (b=Black, h=Hispanic, w=White)
- z1 racial group (1=black, 0=white)
- z2 ethnic group (1=Hispanic, 0=non-Hispanic)

Source

<http://www.stat.ufl.edu/~aa/social/data.html>

Examples

```
data(inc.ed.race.13p1)
summary(inc.ed.race.13p1)
```

income.credit

Income and Credit Card Possession Data

Description

Data on annual income (euros), number of subjects at that income level, and number possessing##' a travel credit card. Source: Based on data in "Categorical Data Analysis," Quaderni del Corso Estivo di Statistica e Calcolo delle Probabilita;, n. 4., Istituto di Metodi Quantitativi, Universita; Luigi Bocconi, a cura di R. Piccarreta (1993). Table 15.1 of 4th edition.

Format

data.frame with 24 observations of the following 3 variables,

- Income annual income (euros)
- n number of subjects
- credit number possessing a credit card

Source

<http://www.stat.ufl.edu/~aa/social/data.html>

Examples

```
data(income.credit)
summary(income.credit)
```

mental.impairment	<i>Mental Impairment Data</i>
-------------------	-------------------------------

Description

Data from a study in Alachua County, Florida on between mental health. Source A. Charles Holzer. See Table 11.1 of the 4th edition.

Format

data.frame with 40 observations of the following 3 variables,
 impair Mental impairment
 life Life events score
 ses Social economic status

Source

<http://www.stat.ufl.edu/~aa/social/data.html>

Examples

```
data(mental.impairment)
summary(mental.impairment)
```

oecd.data	<i>OECD Data</i>
-----------	------------------

Description

The OECD (Organization for Economic Cooperation and Development) is an international organization that consists of developed countries. This dataset has UN data on OECD countries. See Table 3.11 and Exercise 3.6.

Format

data.frame with 23 observations of the following 9 variables,
 nation Country name
 GDP GDP per capita (in US dollars)
 Unemploy Percent unemployed
 Inequal A measure of inequality that compares the wealth of the richest 10% to the poorest 10%
 Health Public expenditure on health (as a percent of GDP)
 Phys Number of physicians per 100,000 people
 CO2 Carbon dioxide emissions (per capita, in metric tons)
 Parlia The percentage of seats in parliament held by women.
 FemEcon Female economic activity as a percentage of the male rate

Source

<http://www.stat.ufl.edu/~aa/social/data.html>

Examples

```
data(oecd.data)
summary(oecd.data)
```

statewide.crime.2	<i>Statewide Crime 2</i>
-------------------	--------------------------

Description

The data are from Statistical Abstract of the United States and most variables were measured in 1993. Table 9.1 of the 3rd edition, and some appear in Table 9.1 of the 4th edition.

Format

data.frame with 51 observations of the following 8 variables,

State U.S. State

VR violent crime rate (per 100,000 people in population)

MR murder rate (per 100,000 people in population)

M percent in metropolitan areas

W percent white

H percent high school graduates

P percent with income below the poverty level

S percent of families headed by a single parent.

Source

<http://www.stat.ufl.edu/~aa/social/data.html>

Examples

```
data(statewide.crime.2)
summary(statewide.crime.2)
```

`student.survey`*Student Survey data*

Description

This data file consists of responses of graduate students in the social sciences enrolled in STA 6126 in a recent term at the University of Florida.

Format

`data.frame` with 60 observations of the following 18 variables,

GE gender

AG age in years

HI high school GPA (on a four-point scale)

CO college GPA

DH distance (in miles) of the campus from your home town

DR distance (in miles) of the classroom from your current residence

TV average number of hours per week that you watch TV

SP average number of hours per week that you participate in sports or have other physical exercise

NE number of times a week you read a newspaper

AH number of people you know who have died from AIDS or who are HIV+

VE whether you are a vegetarian

PA political affiliation (d = Democrat, r = Republican, i = independent)

PI political ideology

RE how often you attend religious services

AB opinion about whether abortion should be legal in the first three months of pregnancy

AA support affirmative action

LD belief in life after death

Source

<http://www.stat.ufl.edu/~aa/social/data.html>

Examples

```
data(student.survey)
summary(student.survey)
```

UNdata

Data from 2005 UN Human Development Report.

Description

These data are from the Human Development report of 2005. Note however that the values given are from 2003 except for CO2 which is 2002.

Table 9.13 in the 4th Edition.

Format

data.frame with 39 observations of the following 11 variables,

HDI HDI value

Fert Total Fertility rate (births/woman)

Cont Contraceptive prevalence rate (%)

Cell Cellular subscribers (per 1000 people)

Inter Internet users (per 1000 people)

GDP GDP per capita (US\$)

CO2 Carbon dioxide emissions per capita (metric tons)

Life Life expectancy at birth, female (years)

Liter Adult literacy rate (female rate % ages 15 and above)

FemEc Female economic activity rate (% of male rate, ages 15 and above)

Source

<http://www.stat.ufl.edu/~aa/social/data.html>

Examples

```
data(UNdata)
summary(UNdata)
```

us.pop.size

U.S. Population Size Data

Description

US population size by decade. Table 14.8 of 4th edition.

Format

data.frame with 12 observations of the following 2 variables,

decade Decade (0=1890, 11=2000)

population Population (millions)

Source

<http://www.stat.ufl.edu/~aa/social/data.html>

Examples

```
data(us.pop.size)
summary(us.pop.size)
```

zagat

Zagat restaunt rating of Italian Restraunts Data

Description

Zagat restaunt rating of Italian Restraunts Data. Examples 3.59, 3.60, and 9.38 of 4th edition.

Format

data.frame with 193 observations of the following 6 variables,

City City

Restaurant Restaurant

Food Food

Decor Decor

Service Service

Cost Cost

Source

<http://www.stat.ufl.edu/~aa/social/data.html>

Examples

```
data(zagat)
summary(zagat)
```

Index

`anorexia`, [2](#)

`birth.rates`, [3](#)

`crime2005`, [3](#)

`fertility.gdp`, [4](#)

`fl.crime`, [5](#)

`house.selling.price`, [5](#)

`house.selling.price.2`, [6](#)

`inc.ed.race.13p1`, [6](#)

`income.credit`, [7](#)

`mental.impairment`, [8](#)

`oecd.data`, [8](#)

`smss-package`, [2](#)

`statewide.crime.2`, [9](#)

`student.survey`, [10](#)

`UNdata`, [11](#)

`us.pop.size`, [12](#)

`zagat`, [12](#)