Package 'GmooG'

September 2, 2024

Type Package

Title Datasets for the Book 'Getting (more out of) Graphics'

Version 0.7

Description Datasets analysed in the book Antony Unwin (2024, ISBN:978-0367674007) ``Getting (more out of) Graphics".

Depends R (>= 3.5)

License GPL (>= 2)

Suggests tidyverse

NeedsCompilation no

Author Antony Unwin [aut, cre, cph]

Maintainer Antony Unwin <unwin@math.uni-augsburg.de>

Repository CRAN

Date/Publication 2024-09-02 15:00:02 UTC

Contents

---

aFacial   *Testing facial recognition software*

Description

Buolamwini and Gebru used their own database that included more women and more people of colour to evaluate how well commercial gender classification algorithms coped with different shades of skin colour in a gender-balanced test database.

Usage

```
data(aFacial)
```

Format

A data frame with 72 observations on the following 5 variables.

Sex Female or Male

Skin one of six shades of skin colour from I to VI

Prediction Correct or Wrong

Freq number of cases

Software one of three facial recognition software packages

Details

Summary data tables of percentages and some numerical totals were provided in the paper and the supplementary material. Assuming the results had to be based on integer numbers of cases it was possible to reconstruct summary raw numbers of the dataset. The dataset is analysed in Chapter 22, "Comparing software for facial recognition".

Source

Buolamwini, Joy, and Timnit Gebru. 2018. "Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification." Proceedings of Machine Learning Research 81: 1-15

Examples

```
data(aFacial, package="GmooG")
head(aFacial, n=12)
```

All200 *The 200 best times for male and female swimmers for many swimming events*

Description

The best times up till mid-2021 are for 17 individual swimming events for men and women and for three relay events.

Usage

```
data(All200)
```

Format

A data frame with 7685 observations on the following 10 variables.

full_name_computed Name of swimmer

team_code country

sdate date of swim

bdate date of birth

SwimTime performance (in seconds)

Gender Women or Men

style one of four swimming strokes or three relay events

distance length of swim with special coding for relays (e.g. 4x100)

dist length of swim in metres

Rank_Order ranking within an event

Details

The dataset is analysed in Chapter 20, "Are swimmers swimming faster?".

Source

https://www.worldaquatics.com/swimming/rankings

Examples

```
data(All200, package="GmooG")
with(All200, table(style))
```

---

astronauts   *Human space flights*

Description

Individuals who travelled into space between 1961 and 2019.

Usage

```
data(astronauts)
```

Format

A data frame with 1277 observations on the following 24 variables.

id id number of record

number id number of individual

nationwide_number national number of individual

name individual's name

original_name name in own language

sex sex of individual

year_of_birth year of birth of individual

nationality nationality

military_civilian military or civilian

selection selection group

year_of_selection selection year

mission_number mission number of individual

total_number_of_missions total missions of individual

occupation role on flight: commander, pilot, flight engineer, ...

year_of_mission Mission year

mission_title Mission name

`ascend_shuttle` Name of ascent shuttle

`in_orbit` Name of spacecraft used in orbit

`descend_shuttle` Name of descent shuttle

`hours_mission` Duration of mission in hours

`total_hrs_sum` Total duration of all missions in hours

`field21` Instances of EVA by mission

`eva_hrs_mission` Duration of extravehicular activities during the mission

`total_eva_hrs` Total duration of all extravehicular activities in hours

Details

This dataset is used in Chapter 10, "Who went up in space for how long?"

Source

<https://github.com/rfordatascience/tidytuesday/tree/master/data/2020/2020-07-14>

Examples

```
data(astronauts, package="GmooG")
library(tidyverse)
nc <- astronauts %>% count(nationality) %>% arrange(-n)
```

DC1912 *Voting at the 1912 Democratic Convention*

Description

The number of votes by each state for each candidate on each ballot for the Democratic nomination for president.

Usage

```
data(DC1912)
```

Format

A data frame with 3939 observations on the following 4 variables.

`State` State or territory name (there were 52)

`Candidate` Name of one of the 13 candidates or 'NotVoting'

`Ballot` Ballot number (1 to 46)

`Votes` Number of votes for the candidate on that ballot from the state

Details

Two other smaller datasets are used in combination with this one for the final plot of Chapter 4 (Figure 4.7), "Voting 46 times to choose a Presidential candidate", the estimated times of the ballots (DC1912ballots) and the adjournment times (DC1912adjourns).

Source

Woodson, Urey. 1912. Official Report of the Proceedings of the Democratic National Convention. Chicago: Peterson linotyping Company

Examples

```
data(DC1912, package="GmooG")
with(DC1912, table(State))
```

DC1912adjourns   *Times of adjournments at the 1912 Democratic Convention*

Description

Times that the six adjournments started and finished, taken from Woodson's convention report.

Usage

```
data(DC1912adjourns)
```

Format

A data frame with 6 observations on the following 2 variables.

StartT Date and time of start of adjournment

EndT Date and time of end of adjournment

Details

This dataset is used in combination with the datasets DC1912 and DC1912ballots for the final plot of Chapter 4 (Figure 4.7), "Voting 46 times to choose a Presidential candidate".

Source

Woodson, Urey. 1912. Official Report of the Proceedings of the Democratic National Convention. Chicago: Peterson linotyping Company

Examples

```
data(DC1912adjourns, package="GmooG")
DC1912adjourns
```

DC1912ballots   *Estimated times of ballots at the 1912 Democratic Convention*

Description

The date and time that each ballot took place have been estimated from Woodson's convention report.

Usage

```
data(DC1912ballots)
```

Format

A data frame with 46 observations on the following 2 variables.

Ballot Ballot number (1 to 46)

DateT Date and time of the ballot

Details

This dataset is used in combination with the datasets DC1912 and DC1912adjourns for the final plot of Chapter 4 (Figure 4.7), "Voting 46 times to choose a Presidential candidate".

Source

Woodson, Urey. 1912. Official Report of the Proceedings of the Democratic National Convention. Chicago: Peterson linotyping Company

Examples

```
data(DC1912ballots, package="GmooG")
head(DC1912ballots)
```

DC1912dels   *Numbers of delegates for the individual states and groups*

Description

The number of pledged delegates by group at the 2020 Democratic convention.

Usage

```
data(DC1912dels)
```

Format

A data frame with 58 observations on the following 3 variables.

State Name of group (mostly state or territory)

TotP Number of pledged delegates by group at the 2020 Democratic convention

region Ordered factor: MidWest, NorthEast, West, South, Territory, NA

Details

This dataset is used in Chapter 4, "Voting 46 times to choose a Presidential candidate".

Source

https://ballotpedia.org/Democratic_delegate_rules,_2020 and https://www.census.gov

Examples

```
data(DC1912dels, package="GmooG")
head(DC1912dels)
```

DC1912evs   *Electoral votes for the individual states of the US*

Description

The number of electoral votes for each of the 50 states and D.C. from 1788 till 2020.

Usage

```
data(DC1912evs)
```

Format

A data frame with 51 observations on the following 36 variables.

Code Code for State

State State name (there were 51 including D.C.)

y1788 Numbers of electoral votes by State in 1788

y1792 Numbers of electoral votes by State in 1792

y17961800 Numbers of electoral votes by State for 1796 and 1800

y18041808 Numbers of electoral votes by State in 1804 and 1808

y1812 Numbers of electoral votes by State in 1812

y1816 Numbers of electoral votes by State in 1816

y1820 Numbers of electoral votes by State in 1820

y18241828 Numbers of electoral votes by State in 1824 and 1828

y1832 Numbers of electoral votes by State in 1832

y18361840 Numbers of electoral votes by State in 1836 and 1840

y1844 Numbers of electoral votes by State in 1844

y1848 Numbers of electoral votes by State in 1848

y18521856 Numbers of electoral votes by State in 1852 and 1856

y1860 Numbers of electoral votes by State in 1860

y1864 Numbers of electoral votes by State in 1864

y1868 Numbers of electoral votes by State in 1868

y1872 Numbers of electoral votes by State in 1872

y18761880 Numbers of electoral votes by State in 1876 and 1880

y18841888 Numbers of electoral votes by State in 1884 and 1888

y1892 Numbers of electoral votes by State in 1892

y18961900 Numbers of electoral votes by State in 1896 and 1900

y1904 Numbers of electoral votes by State in 1904

y1908 Numbers of electoral votes by State in 1908

y19121928 Numbers of electoral votes by State from 1912 to 1928

y19321940 Numbers of electoral votes by State from 1932 to 1940

y19441948 Numbers of electoral votes by State in 1944 and 1948

y19521956 Numbers of electoral votes by State in 1952 and 1956

y1960 Numbers of electoral votes by State in 1960

y19641968 Numbers of electoral votes by State in 1964 and 1968

y19721980 Numbers of electoral votes by State from 1972 to 1980

y19841988 Numbers of electoral votes by State in 1984 and 1988

y19922000 Numbers of electoral votes by State from 1992 to 2000

y20042008 Numbers of electoral votes by State in 2000 and 2008

y20122020 Numbers of electoral votes by State from 2012 to 2020

Details

This dataset is used in Chapter 4, "Voting 46 times to choose a Presidential candidate".

Source

https://en.wikipedia.org/wiki/United_States_Electoral_College

Examples

```
data(DC1912evs, package="GmooG")
head(DC1912evs[, c("State", "y1788", "y19121928", "y20122020")])
```

Decath21   *The top 116 decathletes of recent times in April 2021*

Description

Details of the best performances of the top decathletes

Usage

```
data(Decath21)
```

Format

A data frame with 116 observations on the following 15 variables.

Rank Rank order

Decathlete Decathlete's name

Nationality Decathlete's nationality

Total the total points achieved over all 10 events

Run100m Time for the 100 metres (secs)

LongJump Distance jumped (metres)

ShotPut Distance putting the shot (metres)

HighJump Height jumped (metres)

Run400m Time for the 400 metres (secs)

Hurdle110m Time for the 110 metres hurdles (secs)

DiscusD Distance throwing the discus (metres)

PoleVault Height achieved (metres)

JavelinD Distance throwing the javelin (metres)

Run1500m Time for the 1500 metres (secs)

Venue Location and year of performance

Source

<https://www.decathlon2000.com>

Examples

```
data(Decath21, package="GmooG")
with(Decath21, summary(Run1500m))
```

DLQI   *DLQI assessment in a phase 3 clinical trial of patients with psoriasis.*

Description

150 psoriasis patients were randomized to Placebo (Treatment A) and 450 to the active treatment (Treatment B). The treatment effect in terms of Quality of Life was assessed at Week 16.

Usage

```
data(DLQI)
```

Format

A data frame with 900 observations on the following 15 variables.

USUBJID individual ID

TRT Placebo (A) or Treatment (B)

PASI_BASELINE Psoriasis Area and Severity Index at Baseline

VISIT Initial or at Week 16

DLQI101 How Itchy, Sore, Painful, Stinging: 0-3

DLQI102 How Embarrassed, Self Conscious: 0-3

DLQI103 Interfered Shopping, Home, Yard: 0-3

DLQI104 Influenced Clothes You Wear: 0-3

DLQI105 Affected Social, Leisure Activity: 0-3

DLQI106 Made It Difficult to Do Any Sports: 0-3

DLQI107 Prevented Working or Studying: 0-3

DLQI108 Problem Partner, Friends, Relative: 0-3

DLQI109 Caused Any Sexual Difficulties: 0-3

DLQI110 How Much a Problem is Treatment: 0-3

DLQI_SCORE DLQI Total Score: 0-30

Details

This dataset is used in Chapter 12, "Psoriasis and the Quality of Life".

Source

<https://github.com/VIS-SIG/Wonderful-Wednesdays/tree/master/data/2021/2021-01-13>

Examples

```
data(DLQI, package="GmooG")
with(DLQI, summary(PASI_BASELINE))
```

DVCdeer   *Vehicle accidents with deer in Bavaria*

Description

Numbers of vehicle accidents with deer every half-hour from the beginning of 2002 till the end of 2011.

Usage

```
data(DVCdeer)
```

Format

A data frame with 175296 observations on the following 3 variables.

mins beginning of half-hour period, from 00:00 to 23:30

day day

Freq number of accidents

Details

This dataset and the dataset DVCnot are both used in Chapter 24, "When do road accidents with deer happen in Bavaria?".

Source

<https://www.jstatsoft.org/article/view/v092i01>

Examples

```
data(DVCdeer, package="GmooG")
with(DVCdeer, table(Freq))
```

DVCnot   *Vehicle accidents in Bavaria not involving deer*

Description

Numbers of vehicle accidents every half-hour from the beginning of 2002 till the end of 2011.

Usage

```
data(DVCnot)
```

Format

A data frame with 175296 observations on the following 3 variables.

`mins` beginning of half-hour period, from 00:00 to 23:30

`day` day, from 2002-01-01 to 2011-12-31

`Freq` number of accidents

Details

This dataset and the dataset DVCnot are both used in Chapter 24, "When do road accidents with deer happen in Bavaria?".

Source

<https://www.jstatsoft.org/article/view/v092i01>

Examples

```
data(DVCnot, package="GmooG")
with(DVCnot, table(Freq))
```

ElecCars   *Trial of how drivers used electric car charging facilities*

Description

A field experiment on electric vehicle charging

Usage

```
data(ElecCars)
```

Format

A data frame with 3395 observations on these 24 variables.

`sessionId` charging session

`kwhTotal` total energy use of a given EV charging session, measured in kWh

`dollars` amount paid by the user in US$ for a given charging session

`created` date and time the session began

`ended` date and time the session ended

`startTime` hour of day began

`endTime` hour of day ended

`chargeTimeHrs` total length of session

`weekday` day of the week of session

`platform` digital platform used by driver

`distance` distance from home, if reported

`userId` user code

`stationId` station code

`locationId` location code

`managerVehicle` binary, 1 if manager car

`facilityType` type of facility, manufacturing = 1, office = 2, research and development = 3, other = 4

`Mon` binary for day of week of session

`Tues` binary for day of week of session

`Wed` binary for day of week of session

`Thurs` binary for day of week of session

`Fri` binary for day of week of session

`Sat` binary for day of week of session

`Sun` binary for day of week of session

`reportedZip` binary, 1 if user reported zip code

Details

This dataset is used in Chapter 13, "Charging electric cars".

Source

doi:10.7910/DVN/NFPQLW

Examples

```
data(ElecCars, package="GmooG")
with(ElecCars, table(weekday))
```

---

eu20col   *Colours worn by European international football teams*

Description

Colours for displaying teams

Usage

```
data(eu20col)
```

Format

A data frame with 39 observations on these 6 variables.

`team_alpha3` three letter short form for country

`url_team` webpage for country

`kit_shirt` shirt colour in hex format

`kit_away` away shirt colour in hex format

`kit_shorts` shorts colour in hex format

`kit_socks` socks colour in hex format

Details

This dataset and the dataset eu20p are both used in Chapter 15, "Home or away: where do soccer players play?"

Source

<https://github.com/guyabel/chord-uefa-ec/>

Examples

```
data(eu20col, package="GmooG")
head(eu20col)
```

---

eu20p	*Colours worn by European international football teams*

Description

Colours for displaying teams

Usage

```
data(eu20p)
```

Format

A data frame with 4012 observations on these 21 variables.

`year` year of competition

`squad` country

`no` player's squad number (from 1968 on)

`pos` position, GK=Goalkeeper, DF=Defender, MF=midfield, FW=Forward

`player` player name

`date_of_birth_age` date of birth and age at competition

caps number of international caps

club club team of player

player_url webpage for player

club_fa_url webpage for Country Football Association of club

club_fa Country Football Association of club

club_2 Second name for club

club_country Country of club

club_country_flag Image of country's flag

goals number of goals scored for country

captain logical TRUE (captain) or FALSE

player_original player name and whether they were captain

nat_team International team

club_country_harm Country of club

nat_team_alpha3 abbreviation for international team

club_alpha3 abbreviation for country of club

Details

This dataset and the dataset eu20col are both used in Chapter 15, "Home or away: where do soccer players play?"

Source

<https://github.com/guyabel/chord-uefa-ec/>

Examples

```
data(eu20p, package="GmooG")
with(eu20p, table(pos))
```

France1954   *Working population of France in 1954*

Description

Numbers working in three sectors in each department of France in 1954.

Usage

```
data(F1954)
```

Format

A data frame with 90 observations on the following 8 variables.

ID ID code for the department

Dept Department name

I.Agriculture Number in thousands of workers in agriculture

II.Industry Number in thousands of workers in industry

III.Commerce Number in thousands of workers in commerce

BertinTotal Total of the three sectors reported by Bertin

Area Area of department in sq kms

NOM_DEPT Alternative name for department

Details

The sector data is from Bertin, while area data has been taken from the Guerry package and Wikipedia. The alternative department name was used for merging with a shape file of France (France54Map). The dataset is analysed in Chapter 7, "Re-viewing Bertin's main example".

Source

Bertin, Jaques. 1973. Semiologie Graphique. 2nd ed. The Hague: Mouton-Gautier

Examples

```
data(F1954, package="GmooG")
with(F1954, summary(I.Agriculture))
```

France54Map   *Map of the departments of France in 1954*

Description

A polygon map of the French departments

Usage

```
data(France54Map)
```

Format

An sf object with 90 observations on the following 2 variables

Dept Department name

geometry list of department polygons

Details

This shape file is used in Chapter 7, "Re-viewing Bertin's main example", and combined with the data in the file F1954. Combining the six new departments of 1967 into the two former departments of Seine and Seine-et-Oise is approximately right.

Source

<http://coulmont.com/cartes/rcarto.pdf> Derived from GEOFLADept_FR_Corse_AV_L93/DEPARTEMENT.SHP

GapLifeE	*Life expectancy data from Gapminder*

Description

Life expectancy at birth for almost 200 countries from 1800 to 2016 and forecasts for 2017 to 2100

Usage

```
data(GapLifeE)
```

Format

A data frame with 187 observations on 302 variables. The first variable is the name of the country. Every other variable is named as a year from 1800 to 2100 and the values are the historical life expectancy figures up to 2016 and forecasts of life expectancy from 2017 on.

Details

This dataset and the datasets GapRegions and GapPop are all used in Chapter 2, "Graphics and Gapminder".

Source

<https://www.gapminder.org>

Examples

```
data(GapLifeE, package="GmooG")
library(tidyverse)
ggplot(GapLifeE, aes(`1900`, `2000`)) + geom_point()
```

GapPop   *Population data from Gapminder*

Description

Population data for almost 200 countries from 1800 to 2016 and forecasts for 2017 to 2100

Usage

```
data(GapPop)
```

Format

A data frame with 195 observations on 302 variables. The first variable is the name of the country. Every other variable is named as a year from 1800 to 2100 and the values are the historical population figures up to 2016 and forecasts of population from 2017 on.

Details

This dataset and the datasets GapLifeE and GapRegions are all used in Chapter 2, "Graphics and Gapminder".

Source

<https://www.gapminder.org>

Examples

```
data(GapPop, package="GmooG")
library(tidyverse)
ggplot(GapPop, aes(`1900`, `2000`)) + geom_point()
```

GapRegions   *World region definitions used by Gapminder*

Description

Gapminder offers several different divisions into regions of the almost 200 countries of the world.

Usage

```
data(GapRegions)
```

Format

A data frame with 197 observations on 16 variables.

geo country abbreviation

name country name

four_regions world split into four regions

eight_regions world split into eight regions

six_regions world split into six regions

members_oecd_g77 group membership: oecd, g77, other

Latitude latitude of country

Longitude longitude of country

UN member since date of joining UN

World bank region world split into seven regions by World bank

World bank, 4 income groups 2017 world split into four income groups by World bank

World bank, 3 income groups 2017 world split into three income groups by World bank, all NA

Details

This dataset and the datasets GapLifeE and GapPop are all used in Chapter 2, "Graphics and Gap-minder".

Source

https://www.gapminder.org

Examples

```
data(GapRegions, package="GmooG")
with(GapRegions, table(four_regions, six_regions))
```

GermanDemographics *Demographic and economic data for Germany in 2021*

Description

Demographic and cconomic data for the 299 German parliamentary constituencies in 2021

Usage

```
data(GermanDemographics)
```

Format

A data frame with 299 observations on the following 17 variables

WkrNr Constituency (Wahlkreis) number

WkrName Constituency name

Communities Number of communities

Area Area in square kms

Population Population

Germans Number of Germans in the population

Foreigners Percentage of foreigners in the population

PopDensity Population density, numbers per square km

Under18 Percentage population under 18

Age1824 Percentage population between 18 and 24

Age2534 Percentage population between 25 and 34

Age3559 Percentage population between 35 and 59

Age6074 Percentage population between 60 and 74

Age75up Percentage population 75 and older

CarsPerP Cars per 1000 people

Hochschulreife Percentage qualified for university

Unemployed Unemployment rate

Details

This dataset and the datasets GermanElection21 and GermanExtraSeats are all used in Chapter 26, "German Election 2021–what happened?"

Source

<https://www.bundeswahlleiterin.de> Derived from btw21_strukturdaten.csv

Examples

```
data(GermanDemographics, package="GmooG")
with(GermanDemographics, summary(Under18))
```

GermanElection21   *Results of the election for the German Bundestag in Autumn 2021*

Description

Detailed results by constituency for the German election of 2021 (and for the previous election in 2017)

Usage

```
data(GermanElection21)
```

Format

A data frame with 16024 observations on the following 9 variables

WkNr Constituency (Wahlkreis) number

WkName Constituency name

Land Bundesland number

Partei Party

Stimme First (personal) or second (party) vote

Anzahl Number of votes in 2021 election

VorpAnzahl Number of votes in 2017 election

Bundesland Bundesland name

Region Region: West, Berlin, East

Details

This dataset and the datasets GermanDemographics and GermanExtraSeats are all used in Chapter 26, "German Election 2021–what happened?"

Source

<https://www.bundeswahlleiterin.de> Derived from btw21_kerg2.csv

Examples

```
library(tidyverse)
data(GermanElection21, package="GmooG")
btw1vP <- GermanElection21 %>% count(Partei) %>% arrange(-n)
```

GermanExtraSeats *Extra seats at German elections from 1949 to 2021*

Description

Numbers of extra seats (Ueberhangmandate and Ausgleichsmandate) needed to satisfy the German election rules

Usage

```
data(GermanExtraSeats)
```

Format

A data frame with 20 observations on these 2 variables.

Year Election year

Number Number of extra seats needed

Details

This dataset is used in Chapter 26, "German Election 2021–what happened?".

Source

German election results from https://www.bundeswahlleiter.de

Examples

```
data(GermanExtraSeats, package="GmooG")
library(tidyverse)
ggplot(GermanExtraSeats, aes(Year, Number)) + geom_line()
```

GermanyMap *Map of the German parliamentary constituencies in 2021*

Description

A polygon map of the German constituencies

Usage

```
data(GermanyMap)
```

Format

An sf object with 299 observations on the following 5 variables

`WKR_NR` Constituency (Wahlkreis) number

`WKR_NAME` Constituency name

`LAND_NR` Bundesland number

`LAND_NAME` Bundesland name

`geometry` list of constituency polygons

Details

This map file is used in Chapter 26, "German Election 2021–what happened?"

Source

<https://www.bundeswahlleiterin.de> Derived from Geometrie_Wahlkreise_20DBT_geo.shp

---

GmooG	*GmooG: datasets analysed in "Getting (more out of) Graphics"*

Description

There are 25 chapters of graphical data analyses in the book. Datasets that are not readily available are mainly provided in this package.

Details

Other datasets are analysed in the book as well. They are available in various R packages. Some can be downloaded and updated from the web.

Author(s)

Antony Unwin unwin@math.uni-augsburg.de

| malaria | *Comparison of four tests for malaria* |

Description

Studying magneto-optical diagnosis of symptomatic malaria in Papua New Guinea.

Usage

```
data(malaria)
```

Format

A data frame with 956 observations on the following 24 variables.

ID Patient ID

Collect_Date Date blood sample collected

Age Patient age

Weight Patient weight

Sex Patient sex

Temperature ancillary temperature in degrees Centigrade

Hb Patient hemoglobin level in g/dL

illMalaria Malaria in last two weeks

RDT1 HRP2 line positive

RDT2 LDH line positive

RDTb HRP and LDH lines positive

Pf qPCR copy number for P. falciparum per microL of blood

Pv qPCR copy number for P. vivax in copies per microL of blood

LM_Pf final expert light microscopy result for P. falciparum in parasites per microL of blood

LM_Pfg final expert light microscopy result for P. falciparum gametocytes in parasites per microL of blood

LM_Pv final expert light microscopy result for P. vivax in parasites per microL of blood

LM_Pvg final expert light microscopy result for P. vivax gametocytes in parasites per microL of blood

LM_Pm final expert light microscopy result for P. malariae in parasites per microL of blood

LM_Po final expert light microscopy result for P. ovale in parasites per microL of blood

AveMO Average magneto-optical signalof blood aliquots #1,2,3 in mV/V

sdMO Standard deviation of the magneto-optical signals of blood aliquots #1,2,3 in mV/V

MO1 Magneto-optical signal of blood aliquot #1 in mV/V

MO2 Magneto-optical signal of blood aliquot #2 in mV/V

MO3 Magneto-optical signal of blood aliquot #3 in mV/V

Details

This dataset is used in Chapter 19, "Comparing tests for malaria".

Source

doi:10.6084/m9.figshare.13078181.v1

Examples

```
data(malaria, package="GmooG")
with(malaria, summary(AveMO))
```

Mich1879	*Measurements of the speed of light by Michelson in 1879*

Description

Michelson included more details of each experiment in the table of results in his report.

Usage

```
data(Mich1879)
```

Format

A data frame with 100 observations on the following 4 variables.

Date Day of the experiment (from 5 June to 2 July 1879)

Time AM, PM or Elec (under electric light)

Value estimate of the speed of light minus 299000, uncorrected for temperature and refraction

Temperature temperature in degrees Fahrenheit, from 58 to 90

Details

This dataset and the dataset newcomb are both used in Chapter 5, "Measuring the speed of light".

Source

Michelson, Albert. 1880. "Experimental Determination of the Velocity of Light Made at the U.S. Naval Academy, Annapolis." Astronomical Papers 1: 109-45. https://books.google.de/books?id=343nAAAAMAAJ

Examples

```
data(Mich1879, package="GmooG")
with(Mich1879, summary(Temperature))
```

---

newcomb   *Measurements of the speed of light by Newcomb in 1882*

Description

Newcomb reported three series of measurements and regarded the third series used here as the best.

Usage

```
data(newcomb)
```

Format

A data frame with 66 observations on the following 6 variables.

Date Day of the experiment (from 24 July to 5 September 1882)

Observer Newcomb or Holcombe (who assisted Newcombe in these experiments)

Wt1 a weight given by Newcomb for the quality of the image observed

Wt2 a second weight for the quality of the image

Time time taken in millionths of a second for light to travel a distance of 7.44242 kilometres in air

Wt overall weight given by Newcomb to the observation

Details

This dataset and the dataset Mich1879 are both used in Chapter 5, "Measuring the speed of light".

Source

Newcomb, Simon. 1891. "Measures of the Velocity of Light Made Under the Direction of the Secretary of the Navy During the Years 1880-1882." Astronomical Papers 2: 107-230

Examples

```
data(newcomb, package="GmooG")
with(newcomb, summary(Time))
```

`OlympicPeople` *Competitors at the modern Olympic Games*

Description

Individuals who competed at the Olympic Games from 1896 to 2016.

Usage

```
data(OlympicPeople)
```

Format

A data frame with 219434 observations on the following 4 variables.

Sex Sex of athlete

NOC Abbreviation for national team

Year Year of Games

City Location of Games

Details

This dataset and the dataset OlympicPerfs are both used in Chapter 6, "The modern Olympic Games in numbers".

Source

Derived from https://www.kaggle.com/datasets/heesoo37/120-years-of-olympic-history-athletes-and-resul

Examples

```
data(OlympicPeople, package="GmooG")
with(OlympicPeople, table(Year))
```

`OlympicPerfs` *Performances of competitors at the modern Summer Olympic Games*

Description

Performances at the Summer Olympic Games from 1896 to 2016.

Usage

```
data(OlympicPerfs)
```

Format

A data frame with 108789 observations on the following 8 variables.

rank rank in event

medalType medal won: one of Gold, Silver, Bronze, NA

games location and year

discipline discipline of event

event name of event

result_value result reported

result_type type of result: distance, time, points, weight, and four others

country country

Details

This dataset and the dataset OlympicPeople are both used in Chapter 6, "The modern Olympic Games in numbers".

Source

Derived from a dataset scraped from the web and provided to the maintainer.

Examples

```
data(OlympicPerfs, package="GmooG")
library(tidyverse)
OlyD <- OlympicPerfs %>% count(discipline)
```

SeaBirds	*Descriptions of three species of shearwaters (Audubon, Galapagos, Tropical)*

Description

Plumage and morphological characteristics of three species of shearwaters.

Usage

```
data(SeaBirds)
```

Format

A data frame with 153 observations on the following 6 variables.

collar one of five categories

eyebrows four levels from none to very pronounced

undertail four levels: White, Black, Black & White, Black & WHITE

border none, few or many

sex male or female

species one of Audubon, Galapagos, Tropical

Details

This dataset is used in Chapter 23, "Distinguishing shearwaters".

Source

Derived from the R package CoModes (numerial categories have been converted to text and common names rather than scientific names are used for species)

Examples

```
data(SeaBirds, package="GmooG")
with(SeaBirds, table(species))
```

SurvGR   *Responses on gay rights in Annenberg's 2004 National Election survey*

Description

Responses on questions about gay rights at State level and Federal level

Usage

```
data(SurvGR)
```

Format

A data frame with 81422 observations on 11 variables.

ID ID number

cDATE Date of interview

State Respondent's state of residence

age Respondent's age

gender Respondent's gender

race Respondent's race

urbanity Urban, Suburban, or Rural

QuF Question answered about Federal gay rights

valF Answer to Federal question

valS Answer to State question

QuS Question answered about State gay rights

Details

This dataset is used in Chapter 9, "Results from surveys on gay rights".

Source

The Annenberg Public Policy Center of the University of Pennsylvania

Examples

```
data(SurvGR, package="GmooG")
with(SurvGR, table(urbanity))
```

TitanicPassCrew   *Passengers and crew who sailed on the Titanic*

Description

Some information on those who sailed on the Titanic

Usage

```
data(TitanicPassCrew)
```

Format

A data frame with 2208 observations on 7 variables.

Age Age of individual

Gender Gender of individual

Group Class of passenger or section of crew

Area abbreviated version of Group

Joined Port where individual boarded:Belfast, Southampton, Cherbourg or Queenstown

Nationality Individual's nationality

survived Whether the individual survived:yes or no

Details

This dataset is used in Chapter 26, "The Titanic Disaster".

Source

Derived from a fuller dataset available from Encyclopedia Titanica

Examples

```
data(TitanicPassCrew, package="GmooG")
with(TitanicPassCrew, table(Joined))
```

USregions   *Map of the Regional Classification of the contiguous US States*

Description

Map of the contiguous US States including information on the regional classification by the Census Bureau

Usage

```
data(USregions)
```

Format

A data frame with 49 observations on 4 variables.

NAME name of state

State 2-letter code for state

Region one of four Census Bureau regions: NorthEast, South, MidWest, West

geometry map polygons for state

Details

This dataset is used in Chapter 9, "Results from surveys on gay rights".

Source

The polygon map data is from the spData package

Examples

```
data(USregions, package="GmooG")
```

VehEffUS	*Fuel economy data for car models in the US*

Description

Fuel economy data for individual models of cars and trucks provided by the US Department of Energy.

Usage

```
data(VehEffUS)
```

Format

A data frame with 43516 observations on the following 16 variables.

year model year, from 1984 to 2022)

make make of car

model model of car

VClass class of vehicle

cylinders number of cylinders, from 2 to 16

atvType type of alternative fuel or advanced technology vehicle

displ engine displacement in liters

drive drive axle type

trany transmission

city city MPG for fuelType1

highway highway MPG for fuelType1

combined combined MPG for fuelType1

fuelCostA08 annual fuel cost for fuelType1 ($)

fuelType1 main fuel type

barrels08 annual petroleum consumption in barrels for fuelType1

co2TailpipeGpm tailpipe CO2 in grams/mile for fuelType1

Details

This dataset is used in Chapter 17, "Fuel efficiency of cars in the USA".

Source

Selection of variables from https://www.fueleconomy.gov/feg/epadata/vehicles.csv.zip

Examples

```
data(VehEffUS, package="GmooG")
with(VehEffUS, table(drive))
```

Index