

# Package ‘gcookbook’

February 19, 2015

**Type** Package

**Title** Data for ‘R Graphics Cookbook’

**Version** 1.0

**Author** Winston Chang

**Maintainer** Winston Chang <winston@stdout.org>

**Description** This package contains data sets used in the book ‘R Graphics Cookbook’ by Winston Chang, published by O’Reilly Media.

**License** GPL-2

**LazyData** yes

**Depends** R (>= 2.10)

**Suggests** ggplot2, plyr, MASS

**Collate** ‘gcookbook.r’

**Repository** CRAN

**Date/Publication** 2012-11-30 12:00:19

**NeedsCompilation** no

## R topics documented:

aapl . . . . .	2
anthoming . . . . .	2
cabbage_exp . . . . .	3
climate . . . . .	3
corneas . . . . .	4
countries . . . . .	4
drunk . . . . .	5
gcookbook . . . . .	5
heightweight . . . . .	5
isabel . . . . .	6
madmen . . . . .	6
madmen2 . . . . .	7

marathon . . . . .	7
pg_mean . . . . .	8
plum . . . . .	8
plum_wide . . . . .	9
simpledat . . . . .	9
simpledat_long . . . . .	10
tophitters2001 . . . . .	10
uspopage . . . . .	11
uspopchange . . . . .	12
wind . . . . .	12
worldpop . . . . .	13

## Index 14

---

aapl	<i>Apple stock data</i>
------	-------------------------

---

### Description

Weekly stock data for AAPL (Apple, Inc.) from 1984 to 2012.

### Variables

- date
- adj\_price: Price, adjusted for splits and dividends.

---

anthoming	<i>Homing in desert ants</i>
-----------	------------------------------

---

### Description

Data from an experiment on the homing performance of a desert ant, *Cataglyphis bicolor*.

### Variables

- angle: Angle between true home direction and the direction that the ant went in (positive is clockwise).
- expt: Number of ants in the experimental condition that went in this direction.
- ctrl: Number of ants in the control condition that went in this direction.

### Source

Hand, D.J., Daly, F., Lunn, A.D., McConway, K.J. & Ostrowski, E. (1994), *A Handbook of Small Data Sets*, Chapman & Hall.

Duelli, P. and Wehner, R. (1973), The spectral sensitivity of polarized light orientation in *Cataglyphis bicolor* (Formicidae, Hymenoptera). *Journal of Comparative Physiology*, **86**, 36-53.

---

cabbage\_exp

*Summary of cabbages data set*

---

### Description

This data set has groupwise means, standard deviations, counts, and standard error of the mean for the `cabbages` data set from the MASS package. The purpose of this summarized data set is to make it easy to use for example graphs.

### See Also

The source data set in the MASS package, `cabbages`.

---

climate

*Global climate temperature anomaly data from 1800 to 2011*

---

### Description

This data set includes estimated global temperature anomaly data for the years 1800 through 2011. The anomaly is the difference from the baseline temperature, which is the mean of the yearly temperatures from 1951-1980.

### Variables

- Source: Data source (Berkeley, CRUTEM3, NASA).
- Year: Year for the estimate.
- Anomaly1y: Temperature anomaly in Celcius, smoothed over one year.
- Anomaly5y: Temperature anomaly in Celcius, smoothed over five years.
- Anomaly10y: Temperature anomaly in Celcius, smoothed over ten years.
- Unc10y: Uncertainty for 10-year-smoothed anomaly.

### Source

Berkeley Earth Project: <http://berkeleyearth.org/dataset/>

Climatic Research Unit (CRUTEM3): <http://www.cru.uea.ac.uk/cru/data/temperature/>

NASA: <http://data.giss.nasa.gov/gistemp/>

---

corneas

*Corneal thickness of eyes*

---

### Description

Corneal thickness of eight people who had glaucoma in one eye.

### Variables

- affected Corneal thickness (in microns) of eye affected by glaucoma.
- notaffected Corneal thickness (in microns) of eye not affected by glaucoma.

### Source

Hand, D.J., Daly, F., Lunn, A.D., McConway, K.J. & Ostrowski, E. (1994), *A Handbook of Small Data Sets*, Chapman & Hall.

Ehlers, N. On corneal thickness and introcular pressure, II. (1970). *Acta Ophthalmologica*, **48**, 1107-1112.

---

countries

*Health and economic data about countries around the world from 1960-2010*

---

### Description

Health and economic data about countries around the world from 1960-2010, from the World Bank.

### Variables

- Name: Name of country
- Code: Short country code
- Year
- GDP: Per capita Gross Domestic Product, in adjusted 2011 U.S. Dollars
- laborrate: Labor rate.
- healthexp: Health expenditures in U.S. Dollars.
- infmortality: Infant mortality per 1000 live births.

### Source

World Bank: <http://data.worldbank.org/>

---

drunk	<i>Convictions for drunkenness</i>
-------	------------------------------------

---

**Description**

Number of people convicted for drunkenness at Tower Bridge and Lambeth Magistrates' Courts from January 1 to June 27, 1970, classified by age and sex.

**Source**

Hand, D.J., Daly, F., Lunn, A.D., McConway, K.J. & Ostrowski, E. (1994), *A Handbook of Small Data Sets*, Chapman & Hall.

Cook, T. (1971). *New Society*, 20 May, 1971.

---

gcookbook	<i>gcookbook: Data sets for "R Graphics Cookbook"</i>
-----------	---

---

**Description**

This package contains data sets used in the book "R Graphics Cookbook" by Winston Chang, published by O'Reilly Media.

---

heightweight	<i>Height and weight of schoolchildren</i>
--------------	--

---

**Description**

Height and weight of schoolchildren

**Variables**

- sex
- ageYear: Age in years.
- ageMonth: Age in months.
- heightIn: Height in inches.
- weightLb: Weight in pounds.

**Source**

Lewis, T., & Taylor, L.R. (1967), *Introduction to Experimental Ecology*, Academic Press.

---

isabel

*Data from simulation of hurricane Isabel*

---

### Description

This data is from a simulation of hurricane Isabel in 2003. It includes temperature and wind data for a 2139km (east-west) x 2004km (north-south) x 19.8km (vertical) volume. The simulation data is from the National Center for Atmospheric Research, and it was used in the IEEE Visualization 2004 Contest.

### Variables

- x: Latitude (x coordinate).
- y: Longitude (y coordinate).
- z: Height in km (z coordinate).
- vx: x wind component in m/s
- vy: y wind component in m/s
- vz: z wind component in m/s
- t: Temperature in Celcius
- speed: wind speed,  $\sqrt{vx^2 + vy^2 + vz^2}$

### Source

<http://vis.computer.org/vis2004contest/data.html>

<http://ncar.ucar.edu/>

---

madmen

*Successful sexual relations in Mad Men (TV show)*

---

### Description

Each row of this data frame represents a pair of characters who had a sexual relationship on the TV show Mad Men, as of the end of season 4. This data can be displayed with an undirected graph.

### Variables

The placement of names in column Name1 as opposed to Name2 is arbitrary, and not meaningful. In other words, for any row, you can swap the values of Name1 and Name2, and it will represent the same information.

- Name1: Name of one sexual partner.
- Name2: Name of another sexual partner.

**Source**

Wired Magazine 20.02, February 2012

**See Also**

For a list of attempted sexual pairings, see [madmen2](#).

---

madmen2

*Attempted sexual relations in Mad Men (TV show)*

---

**Description**

Each row of this data frame represents a pair of characters on the TV show Mad Men, as of the end of season 4. Each row represents an attempted sexual relation: the character in the first column, Name1, attempted to have sex with the character in the second column, Name2. If the relationship goes in both directions (the characters had sex with each other), then there will be two rows, representing each direction. This data can be displayed with a directed graph.

**Variables**

- Name1: Character who made sexual advances.
- Name2: Character who was the target of sexual advances.

**Source**

Wired Magazine 20.02, February 2012

**See Also**

For a list of successful sexual pairings, see [madmen](#).

---

marathon

*Marathon and half-marathon times*

---

**Description**

This data set contains mrathon and half-marathon running times for 520 people. Each row represents one person's times.

**Variables**

- Half: Time in minutes, for half marathon.
- Full: Time in minutes, for full marathon.

**Source**

Downey, A.B. (2011), *Think Stats*, O'Reilly Media.

---

pg\_mean

*Means of results from an experiment on plant growth*

---

### Description

This data set simply has groupwise means of the [PlantGrowth](#) data set. The purpose of this summarized data set is to make it easy to use for example graphs.

### Source

Dobson, A. J. (1983) *An Introduction to Statistical Modelling*, Chapman & Hall.

### See Also

The source data set, [PlantGrowth](#).

---

plum

*Plum root cuttings (long format)*

---

### Description

This is data from an experiment to investigate the effect of cutting length and planting time on the survival of plum root cuttings.

### Variables

- length: Cutting length.
- time: Planting time.
- survival: Survival status.
- count: Number of plants.

### Source

Hand, D.J., Daly, F., Lunn, A.D., McConway, K.J. & Ostrowski, E. (1994), *A Handbook of Small Data Sets*, Chapman & Hall.

Bartlett, M.S. (1935), Contingency table interactions, *Journal of the Royal Statistical Society Supplement*, **2**, 248-252.

### See Also

This data frame is in "long" format. See [plum\\_wide](#) for the same data in "wide" format.



---

plum_wide	<i>Plum root cuttings (wide format)</i>
-----------	---

---

**Description**

This is data from an experiment to investigate the effect of cutting length and planting time on the survival of plum root cuttings.

**Variables**

- length: Cutting length.
- time: Planting time.
- dead: Number of dead plants in this condition.
- alive: Number of alive plants in this condition.

**Source**

Hand, D.J., Daly, F., Lunn, A.D., McConway, K.J. & Ostrowski, E. (1994), *A Handbook of Small Data Sets*, Chapman & Hall.

Bartlett, M.S. (1935), Contingency table interactions, *Journal of the Royal Statistical Society Supplement*, **2**, 248-252.

**See Also**

This data frame is in "wide" format. See [plum](#) for the same data in "long" format.

---

simpledat	<i>Simple example data set</i>
-----------	--------------------------------

---

**Description**

This data set is for examples of R graphics.

**See Also**

This data frame is in "wide" format. See [simpledat\\_long](#) for the same data in "long" format.

---

simpledat_long	<i>Simple example data set (long format)</i>
----------------	--

---

**Description**

This data set is for examples of R graphics.

**See Also**

This data frame is in "long" format. See [simpledat](#) for the same data in "wide" format.

---

tophitters2001	<i>Batting averages of the top hitters in Major League Baseball in 2001</i>
----------------	---

---

**Description**

Batting statistics for the top 144 hitters in Major League Baseball in 2001.

**Variables**

Variables:

- id: Unique player id
- first: First name
- last: Last name
- name: Full name (first and last)
- year: Year of data
- stint
- team: Abbreviation of team played for
- lg: League (American League or National League)
- g: Number of games
- ab: Number of times at bat
- r: Number of runs
- h: Number of hits (times reached base because of a batted, fair ball without error by the defense)
- 2b: Hits on which the batter reached second base safely
- 3b: Hits on which the batter reached third base safely
- hr: Number of home runs
- rbi: Runs batted in
- sb: Stolen bases

- cs: Caught stealing
- bb: Base on balls (walk)
- so: Strike outs
- ibb: Intentional base on balls
- hbp: Hits by pitch
- sh: Sacrifice hits
- sf: Sacrifice flies
- gidp: Ground into double-play
- avg: Batting average (hits divided by at-bats)

**Source**

<http://www.baseball-databank.org/>.

---

uspopage

*Age distribution of population in the United States, 1900-2002*

---

**Description**

These are the estimated (not counted) values by the U.S. Census.

**Variables**

- Year
- AgeGroup
- Thousands: Number of people, in thousands.

**Source**

U.S. Census Bureau, Statistical Abstract of the United States, 2003, HS-3: <http://www.census.gov/statab/hist/HS-03.pdf> and [www.census.gov/statab/hist/02HS0003.xls](http://www.census.gov/statab/hist/02HS0003.xls).

---

 uspopchange

*Change in population of states in the U.S. between 2000 and 2010*


---

### Description

This data set represents the percent change in population of states in the U.S. from 2000 to 2010.

### Variables

- State
- Abb: Abbreviated state name.
- Region: Region of country that the state is in.
- Change: Percent change in population.

### Source

U.S. Census Bureau, Statistical Abstract of the United States, 2012, Table 14. <http://www.census.gov/compendia/statab/2012/tables/12s0014.pdf>

---

 wind

*Wind speed and direction*


---

### Description

This data set contains the wind speed and direction over the course of a single day in Chicago (February 29, 2012).

### Variables

- TimeUTC: Time of day in minutes; 0 is midnight.
- Temp: Temperature in Celcius.
- WindAvg: Average wind speed in m/s in this time block.
- WindMax: Maximum wind speed in m/s in this time block.
- WindDir: Average direction that wind comes from (0=north, 90=east).
- SpeedCat: Average wind speed, categorized in 5 m/s groups.
- DirCat: Average direction, categorized in 15-degree groups.

### Source

Great Lakes Environmental Research Laboratory: <http://www.glerl.noaa.gov/metdata/chi/>

---

worldpop

*World population estimates from 10,000 B.C. to 2,000 A.D.*

---

**Description**

World population estimates from 10,000 B.C. to 2,000 A.D.

**Variables**

- Year
- Population: Estimated population, in thousands

# Index

## \*Topic **datasets**

aapl, 2  
anthoming, 2  
cabbage\_exp, 3  
climate, 3  
corneas, 4  
countries, 4  
drunk, 5  
heightweight, 5  
isabel, 6  
madmen, 6  
madmen2, 7  
marathon, 7  
pg\_mean, 8  
plum, 8  
plum\_wide, 9  
simplifiedat, 9  
simplifiedat\_long, 10  
tophitters2001, 10  
uspopage, 11  
uspopchange, 12  
wind, 12  
worldpop, 13

aapl, 2  
anthoming, 2

cabbage\_exp, 3  
cabbages, 3  
climate, 3  
corneas, 4  
countries, 4

drunk, 5

gcookbook, 5  
gcookbook-package (gcookbook), 5

heightweight, 5

isabel, 6

madmen, 6, 7  
madmen2, 7, 7  
marathon, 7

package-gcookbook (gcookbook), 5  
pg\_mean, 8  
PlantGrowth, 8  
plum, 8, 9  
plum\_wide, 8, 9

simplifiedat, 9, 10  
simplifiedat\_long, 9, 10

tophitters2001, 10

uspopage, 11  
uspopchange, 12

wind, 12  
worldpop, 13