Package ‘dobson’

October 9, 2017

**Type** Package  
**Title** Data from the GLM Book by Dobson and Barnett  
**Version** 0.1  
**License** GPL-2  
**Encoding** UTF-8  
**LazyData** true  
**Depends** R (>= 2.10)  
**RoxygenNote** 6.0.1  
**NeedsCompilation** no  
**Author** Adrian Barnett [aut, cre]  
**Maintainer** Adrian Barnett <a.barnett@qut.edu.au>  
**Repository** CRAN  
**Date/Publication** 2017-10-09 08:44:02 UTC

R topics documented:

- achievement .......................................................... 2  
- aids ................................................................. 3  
- anthers ............................................................ 4  
- balanced .......................................................... 4  
- beetle .............................................................. 5  
- birthweight ......................................................... 6  
- carbohydrate ...................................................... 6  
- Cars ................................................................. 7  
- cholesterol ......................................................... 8  
- chronic ............................................................ 8  
- cyclones ........................................................... 9
<table>
<thead>
<tr>
<th>achievement</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>dobson</td>
<td>9</td>
</tr>
<tr>
<td>doctors</td>
<td>10</td>
</tr>
<tr>
<td>dogs</td>
<td>10</td>
</tr>
<tr>
<td>ear</td>
<td>11</td>
</tr>
<tr>
<td>failure</td>
<td>12</td>
</tr>
<tr>
<td>graduates</td>
<td>12</td>
</tr>
<tr>
<td>hepatitis</td>
<td>13</td>
</tr>
<tr>
<td>hiroshima</td>
<td>14</td>
</tr>
<tr>
<td>housing</td>
<td>14</td>
</tr>
<tr>
<td>insurance</td>
<td>15</td>
</tr>
<tr>
<td>leukemia</td>
<td>16</td>
</tr>
<tr>
<td>machine</td>
<td>16</td>
</tr>
<tr>
<td>melanoma</td>
<td>17</td>
</tr>
<tr>
<td>mortality</td>
<td>17</td>
</tr>
<tr>
<td>moths</td>
<td>18</td>
</tr>
<tr>
<td>pasture</td>
<td>19</td>
</tr>
<tr>
<td>plant.dried</td>
<td>19</td>
</tr>
<tr>
<td>plants</td>
<td>20</td>
</tr>
<tr>
<td>plasma</td>
<td>20</td>
</tr>
<tr>
<td>PLOS</td>
<td>21</td>
</tr>
<tr>
<td>poisson</td>
<td>22</td>
</tr>
<tr>
<td>remission</td>
<td>22</td>
</tr>
<tr>
<td>senility</td>
<td>23</td>
</tr>
<tr>
<td>stroke</td>
<td>23</td>
</tr>
<tr>
<td>sugar</td>
<td>24</td>
</tr>
<tr>
<td>survival</td>
<td>25</td>
</tr>
<tr>
<td>tumor</td>
<td>26</td>
</tr>
<tr>
<td>ulcer</td>
<td>26</td>
</tr>
<tr>
<td>unbalanced</td>
<td>27</td>
</tr>
<tr>
<td>vaccine</td>
<td>28</td>
</tr>
<tr>
<td>waist</td>
<td>28</td>
</tr>
</tbody>
</table>

**Index**

<table>
<thead>
<tr>
<th>achievement</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>achievement</td>
<td>30</td>
</tr>
</tbody>
</table>

**Description**

Achievement scores after three training methods

**Usage**

`data(achievement)`
### Format

A tibble with 21 observations and the following 3 variables.

- **method**: training method (A, B or C)
- **y**: achievement scores
- **x**: aptitude scores measured before training commenced

### References


### Examples

```r
data(achievement)
summary(achievement)
```

### Description

Numbers of cases of AIDS in Australia by date of diagnosis for successive 3-month periods from 1984 to 1988

### Usage

```r
data(aids)
```

### Format

A tibble with 20 observations and the following 3 variables.

- **year**: year
- **quarter**: quarter of year
- **cases**: number of cases

### Source

National Centre for HIV Epidemiology and Clinical Research 1994

### Examples

```r
data(aids)
summary(aids)
```
**Embryogenic anthers data from table 7.2**

**Description**

Numbers of embryogenic anthers of the plant species Datura innoxia Mill obtained when anthers were prepared under several different conditions

**Usage**

data(anthers)

**Format**

A tibble with 6 observations and the following 4 variables.

- `y` numbers of embryogenic anthers
- `n` number of anthers
- `storage` storage condition, control or treatment
- `centrifuge` centrifuging force (g)

**References**


**Examples**

data(anthers)
summary(anthers)

---

**Balanced data from table 6.10**

**Description**

Fictitious balanced data for a two-factor ANOVA with equal numbers of observations in each subgroup

**Usage**

data(balanced)
**beetle**

**Format**

A tibble with 12 observations and the following 3 variables.

factorA factor A
factorB factor B
data dependent data

**Examples**

data(balanced)
summary(balanced)

---

**beetle**

*Beetle data from table 7.2*

---

**Description**

Numbers of beetles dead after five hours exposure to gaseous carbon disulphide at various concentrations

**Usage**

data(beetle)

**Format**

A tibble with 6 observations and the following 3 variables.

x dose (log base 10 CS2mg/l^-1)
n number of beetles
y numbers killed

**References**


**Examples**

data(beetle)
summary(beetle)
### Birthweight

**Birthweight data from table 2.3**

**Description**

Birthweight and gestational age for twelve boys and girls

**Usage**

```r
data(birthweight)
```

**Format**

A tibble with 12 observations and the following 4 variables.

- boys gestational age
- boys gestational age (weeks)
- boys weight
- boys birthweight (grams)
- girls gestational age
- girls gestational age (weeks)
- girls weight
- girls birthweight (grams)

**Examples**

```r
data(birthweight)
summary(birthweight)
```

### Carbohydrate

**Carbohydrate data from table 6.3**

**Description**

Percentages of total calories obtained from complex carbohydrates, for twenty male insulin-dependent diabetics who had been on a high-carbohydrate diet for six months.

**Usage**

```r
data(carbohydrate)
```

**Format**

A tibble with 20 observations and the following 4 variables.

- carbohydrate: percent of total calories obtained from complex carbohydrates
- age: age in years
- weight: body weight relative to "ideal" weight for height
- protein: percentage of calories as protein
Source

K. Webb

Examples

data(carbohydrate)
summary(carbohydrate)

---

Cars  
*Cars data from table 8.1*

Description

Preferences for air conditioning and power steering in cars by gender and age.

Usage

data(Cars)

Format

A tibble with 18 observations and the following 4 variables.

- sex  
- age  
- response  
- frequency

References


Examples

data(Cars)
summary(Cars)
### cholesterol

*Cholesterol data from table 6.21*

**Description**

Cholesterol, age and BMI for thirty women.

**Usage**

```r
data(cholesterol)
```

**Format**

A tibble with 30 observations and the following 3 variables.

- `chol`: serum cholesterol (millimoles per liter)
- `age`: age (years)
- `bmi`: body mass index (kg/m²)

**Examples**

```r
data(cholesterol)
supply(cholesterol)
```

### chronic

*Chronic health data from table 2.7*

**Description**

Numbers of chronic medical conditions reported by samples of women living in large country towns (town group) or in more rural areas (country group) in New South Wales, Australia

**Usage**

```r
data(chronic)
```

**Format**

A data frame with 49 observations and the following 2 variables.

- `place`: place (town or country)
- `number`: number of conditions

**Examples**

```r
data(chronic)
supply(chronic)
```
### cyclones

#### Description

The number of tropical cyclones during a season from November to April in Northeastern Australia.

#### Usage

```r
data(cyclones)
```

#### Format

A tibble with 13 observations and the following 3 variables.

- `years`  season years
- `season` season number
- `number` number of cyclones

#### References


#### Examples

```r
data(cyclones)
summary(cyclones)
```

---

### dobson

#### Description

Datasets from our book *An Introduction to Generalised Linear Models* (4th edition)
**doctors**

*Doctors data from table 9.1*

**Description**

Data from the famous doctors study of smoking conducted by Sir Richard Doll and colleagues.

**Usage**

```r
data(doctors)
```

**Format**

A tibble with 10 observations and the following 4 variables.

- `age` age group
- `smoking` smoker or non-smoker
- `deaths` number of deaths
- `person-years` person years of observation at the time of the analysis

**References**


**Examples**

```r
data(doctors)
summary(doctors)
```

---

**dogs**

*Dogs data from table 11.9*

**Description**

Measurements of left ventricular volume and parallel conductance volume on five dogs under eight different load conditions.

**Usage**

```r
data(dogs)
```
Format

A `tibble` with 40 observations and the following 4 variables.

- `dog` dog number
- `condition` load condition
- `y` left ventricular volume
- `x` parallel conductance volume

References


Examples

```r
data(dogs)
summary(dogs)
```

---

**ear**

Ears data from table 11.10

Description

Numbers of ears clear of acute otitis media at 14 days by antibiotic treatment and age of the child. The children had acute otitis media in both ears.

Usage

```r
data(ear)
```

Format

A `tibble` with 18 observations and the following 4 variables.

- `age` child's age
- `treatment` two treatments coded CEF and AMO
- `number cleared` number of clear ears
- `frequency` faculty

Source


Examples

```r
data(ear)
summary(ear)
```
Failure time data from table 4.1

Description
Lifetimes of Kevlar epoxy strand pressure vessels at 70

Usage
data(failure)

Format
A tibble with 49 observations and the following variable.
lifetimes time to failure in hours

References

Examples
data(failure)
summary(failure)

Graduate survival data from tables 7.14 and 7.15

Description
Survival 50 years after graduation of men and women who graduated each year from 1938 to 1947 from various Faculties of the University of Adelaide.

Usage
data(graduates)

Format
A tibble with 60 observations and the following 5 variables.
year year of graduation
survive number of graduates who survived
total total number of graduates
faculty faculty
sex sex
hepatitis

Source

J.A. Keats

Examples

data(hepatitis)
summary(hepatitis)

hepatitis  Hepatitis data from table 10.5

Description

Survival times in months of patients with chronic active hepatitis in a randomized controlled trial of prednisolone versus no treatment

Usage

data(hepatitis)

Format

A tibble with 44 observations and the following 3 variables.

survival time  survival time in months
censor  censored, lost to follow up or died
group  prednisolone or no treatment

References


Examples

data(hepatitis)
summary(hepatitis)
**hiroshima**  
*Hiroshima data from table 7.12*

**Description**

The number of deaths from leukemia and other cancers among survivors of the Hiroshima atom bomb. The data are for deaths during the period 1950–1959 among survivors who were aged 25 to 64 years in 1950.

**Usage**

```r
data(hiroshima)
```

**Format**

A tibble with 6 observations and the following 4 variables.

- `radiation`: radiation dose (rads)
- `leukemia`: leukemia deaths
- `other_cancer`: deaths from other cancers
- `total_cancers`: total cancer deaths

**References**


**Examples**

```r
data(hiroshima)
summary(hiroshima)
```

---

**housing**  
*Housing data from table 8.5*

**Description**

Data from an investigation into satisfaction with housing conditions in Copenhagen

**Usage**

```r
data(housing)
```
**Format**

A tibble with 18 observations and the following 4 variables.

- **type**: housing type; tower block, apartment or house
- **satisfaction**: satisfaction; low, medium or high
- **contact**: contact with other residents; low or high
- **frequency**: frequency

**References**


**Examples**

```r
data(housing)
summary(housing)
```

---

**insurance**  
*Insurance data from table 9.13*

**Description**

Insurance claim data by car category, age group and district.

**Usage**

```r
data(insurance)
```

**Format**

A tibble with 32 observations and the following 5 variables.

- **car**: car insurance category
- **age**: age group
- **district**: district where policy holder lived; 1=major city, 0=elsewhere
- **y**: number of claims
- **n**: number of insurance policies

**References**


**Examples**

```r
data(insurance)
summary(insurance)
```
**leukemia**

*Leukemia data from table 4.6*

**Description**

Survival times and white blood cell count for seventeen patients suffering from leukemia

**Usage**

```r
data(leukemia)
```

**Format**

A tibble with 17 observations and the following 2 variables.

- `time`: time to death in weeks
- `wbc`: log base 10 initial white blood cell count

**References**


**Examples**

```r
data(leukemia)
summary(leukemia)
```

---

**machine**

*Machine data from table 6.23*

**Description**

Weights of machine components made by workers on different days

**Usage**

```r
data(machine)
```

**Format**

A tibble with 44 observations and the following 3 variables.

- `day`: day number 1 or 2
- `worker`: worker number 1 to 4
- `weight`: weight in grams
Examples

data(melanoma)
summary(melanoma)

melanoma  
*Melanoma data from table 9.4*

Description

A cross-sectional study of patients with a form of skin cancer called malignant melanoma

Usage

data(melanoma)

Format

A tibble with 12 observations and the following 3 variables.

<table>
<thead>
<tr>
<th>variable</th>
<th>description</th>
</tr>
</thead>
<tbody>
<tr>
<td>type</td>
<td>tumor type</td>
</tr>
<tr>
<td>site</td>
<td>site of cancer</td>
</tr>
<tr>
<td>frequency</td>
<td>frequency</td>
</tr>
</tbody>
</table>

References


Examples

data(melanoma)
summary(melanoma)

mortality  
*Mortality data from table 3.2*

Description

Numbers of deaths from coronary heart disease and population sizes by 5-year age groups for men in the Hunter region of New South Wales, Australia in 1991.

Usage

data(mortality)
Format

A tibble with 8 observations and the following 3 variables.

- **age group** age group (years)
- **deaths** number of deaths
- **population** population size

Examples

```r
data(mortality)
summary(mortality)
```

---

**moths**  
*Moths data from table 1.4*

Description

Numbers of females and males in the progeny of 16 female light brown apple moths in Muswellbrook, New South Wales, Australia

Usage

```r
data(moths)
```

Format

A tibble with 16 observations and the following 3 variables.

- **group** progeny group
- **females** number of females
- **males** number of males

References


Examples

```r
data(moths)
summary(moths)
```
**pasture**  

*Pasture data from table 6.20*

**Description**  
Response of a grass and legume pasture system to various quantities of phosphorus fertilizer

**Usage**  
`data(pasture)`

**Format**  
A tibble with 27 observations and the following 2 variables.  
- `k` phosphorus levels (kilograms per hectare)  
- `yield` total yield of grass and legume together (kilograms per hectare)

**Source**  
D. F. Sinclair

**Examples**  
```r  
data(pasture)  
summary(pasture)  
```

---

**plant.dried**  

*Plant data from table 6.7*

**Description**  
Dried weights of plants from three different growing conditions

**Usage**  
`data(plant.dried)`

**Format**  
A tibble with 20 observations and the following 4 variables.  
- `carbohydrate` percent of total calories obtained from complex carbohydrates  
- `age` age in years  
- `weight` body weight relative to "ideal" weight for height  
- `protein` percentage of calories as protein
Source
K. Webb

Examples

data(plant.dried)
summary(plant.dried)

plants  

Description
Plant weight data from table 2.7

Usage
data(plants)

Format
A tibble with 20 observations and the following 2 variables.
treatment  weights of treatment plants in grams
ccontrol  weights of control plants in grams

Examples

data(plants)
summary(plants)

plasma  

Description
Plasma phosphate data from table 6.22

Usage
data(plasma)
Format

A tibble with 31 observations and the following 2 variables.

- **Group**
  - H-O=Hyperinsulinemic obese
  - N-O=Non-hyperinsulinemic obese
  - C=Control

- **phosphate**
  - Plasma inorganic phosphate level (mg/dl)

Examples

data(plasma)
summary(plasma)

---

**PLOS Medicine data from figure 6.3**

Description

Data from 878 journal articles published in PLOS Medicine between 2011 and 2015

Usage

data(PLOS)

Format

A data.frame with 878 observations and the following 2 variables.

- **nchar**
  - Title length

- **authors**
  - Number of authors, truncated to 30

Examples

data(PLOS)
summary(PLOS)
poisson                  Poisson data from table 4.3

Description
Artificial data for a Poisson regression example

Usage
data(poisson)

Format
A tibble with 9 observations and the following two variables.
x covariate
y dependent counts

Examples
data(poisson)
summary(poisson)

remission                Remission data from table 10.1

Description
Times to remission of leukemia patients

Usage
data(remission)

Format
A tibble with 42 observations and the following 3 variables.
time time in weeks
group group; C=control, T=treatment
censored censored; 0=No, 1=Yes

References
**senility**

**Examples**

```r
data(remission)
summary(remission)
```

---

### senility

**Senility data from table 7.8**

**Description**

Data from a sample of elderly people given a psychiatric examination to determine whether symptoms of senility were present together with their score on a subset of the Wechsler Adult Intelligent Scale (WAIS).

**Usage**

```r
data(senility)
```

**Format**

A tibble with 54 observations and the following 2 variables.

- `x` WAIS score
- `s` symptoms of senility present; 1=yes, 0=no

**Examples**

```r
data(senility)
summary(senility)
```

---

**stroke**

**Stroke data from table 11.1**

**Description**

Longitudinal data from an experiment to promote the recovery of stroke patients in wide format. The response variable is the Bartel index with higher scores meaning better outcomes and a maximum score of 100.

**Usage**

```r
data(stroke)
```
**Format**

A tibble with 24 observations and the following 10 variables.

- **Subject**: subject number
- **Group**: group; A = new occupational therapy intervention, B = existing stroke rehabilitation program in the same hospital as A, C = usual care in a different hospital
- **week1**: Bartel index in week 1
- **week2**: Bartel index in week 2
- **week3**: Bartel index in week 3
- **week4**: Bartel index in week 4
- **week5**: Bartel index in week 5
- **week6**: Bartel index in week 6
- **week7**: Bartel index in week 7
- **week8**: Bartel index in week 8

**Source**

C. Cropper, University of Queensland

**Examples**

```r
data(stroke)
summary(stroke)
```

---

**sugar**  
* Sugar data from table 6.19

---

**Description**

Average apparent per capita consumption of sugar (in kg per year) in Australia, as refined sugar and in manufactured foods

**Usage**

```r
data(sugar)
```

**Format**

A tibble with 6 observations and the following 3 variables.

- **period**: period in years
- **refined**: refined sugar
- **manufactured**: Sugar in manufactured food
**Survival data from table 10.1**

**Source**

Australian Bureau of Statistics 1998

**Examples**

```r
data(sugar)
summary(sugar)
```

**Description**

Survival times for leukemia patients

**Usage**

```r
data(survival)
```

**Format**

A tibble with 33 observations and the following 3 variables.

- **survival time**: survival time in weeks
- **WBC**: white blood cell count
- **AG**: test result; + = positive, - = negative

**References**


**Examples**

```r
data(survival)
summary(survival)
```
tumor

**Tumor data from table 8.6**

**Description**

Tumor responses of male and female patients receiving treatment for small-cell lung cancer

**Usage**

```
data(tumor)
```

**Format**

A tibble with 16 observations and the following 4 variables.
- `treatment`: treatment; sequential or alternating
- `sex`: sex
- `response`: four category ordinal response
- `frequency`: frequency

**References**


**Examples**

```
data(tumor)
summary(tumor)
```

ulcer

**Ulcer data from table 9.7**

**Description**

Data from a retrospective case-control study. A group of ulcer patients was compared with a group of control patients not known to have peptic ulcer, but who were similar to the ulcer patients with respect to age, sex and socioeconomic status.

**Usage**

```
data(ulcer)
```
**unbalanced**

**Format**

A tibble with 8 observations and the following 4 variables.

- **ulcer**: type of ulcer
- **case-control**: case or control
- **aspirin**: aspirin user
- **frequency**: frequency

**References**


**Examples**

```r
data(ulcer)
summary(ulcer)
```

**unbalanced**

Unbalanced data from table 6.24

**Description**

Unbalanced data from a fictitious two-factor experiment

**Usage**

```r
data(unbalanced)
```

**Format**

A tibble with 10 observations and the following 3 variables.

- **factorA**: factor A
- **factorB**: factor B
- **data**: dependent data

**Examples**

```r
data(unbalanced)
summary(unbalanced)
```
vaccine

Description
Data from a vaccine trial.

Usage
data(vaccine)

Format
A tibble with 6 observations and the following 3 variables.
treatment treatment group
response response to treatment
frequency frequency

Source
R.S. Gillett

Examples
data(vaccine)
summary(vaccine)

waist

Description
The weights, in kilograms, of twenty men before and after participation in a "waist loss" program

Usage
data(waist)

Format
A tibble with 20 observations and the following 3 variables.
man man number
before weight before in kgs
after weight after in kgs
waist

References


Examples

data(waist)
summary(waist)
Index

*T*opic **datasets**
achievement, 2
aids, 3
anthers, 4
balanced, 4
beetle, 5
birthweight, 6
carbohydrate, 6
Cars, 7
cholesterol, 8
chronic, 8
cyclones, 9
doctors, 10
dogs, 10
ear, 11
failure, 12
graduates, 12
hepatitis, 13
hiroshima, 14
housing, 14
insurance, 15
leukemia, 16
machine, 16
melanoma, 17
mortality, 17
moths, 18
pasture, 19
plant.dried, 19
plants, 20
plasma, 20
PLOS, 21
poisson, 22
remission, 22
senility, 23
stroke, 23
sugar, 24
survival, 25
tumor, 26
ulcer, 26
unbalanced, 27
vaccine, 28
waist, 28
achievement, 2
aids, 3
anthers, 4
balanced, 4
beetle, 5
birthweight, 6
carbohydrate, 6
Cars, 7
cholesterol, 8
chronic, 8
cyclones, 9
dobson, 9
dobson-package (dobson), 9
doctors, 10
dogs, 10
ear, 11
failure, 12
graduates, 12
hepatitis, 13
hiroshima, 14
housing, 14
insurance, 15
leukemia, 16
machine, 16
melanoma, 17
mortality, 17
moths, 18
pasture, 19
plant.dried, 19
plants, 20
plasma, 20
PLOS, 21
poisson, 22
remission, 22
senility, 23
stroke, 23
sugar, 24
survival, 25
tumor, 26
ulcer, 26
INDEX

pasture, 19
plant.dried, 19
plants, 20
plasma, 20
PLOS, 21
poisson, 22
remission, 22
senility, 23
stroke, 23
sugar, 24
survival, 25
tumor, 26
ulcer, 26
unbalanced, 27
vaccine, 28
waist, 28