

# Package ‘Przewodnik’

November 29, 2016

**Title** Datasets and Functions Used in the Book 'Przewodnik po Pakiecie R'

**Version** 0.16.12

**Description** Data sets and functions used in the polish book ``Przewodnik po pakiecie R" (The Hitchhiker's Guide to the R). See more at <<http://biecek.pl/R>>. Among others you will find here data about housing prices, cancer patients, running times and many others.

**Depends** R (>= 3.0.0), PogromcyDanych, PBImisc

**License** GPL-2

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 5.0.1

**NeedsCompilation** no

**Author** Przemyslaw Biecek [aut, cre]

**Maintainer** Przemyslaw Biecek <[przemyslaw.biecek@gmail.com](mailto:przemyslaw.biecek@gmail.com)>

**Repository** CRAN

**Date/Publication** 2016-11-29 12:43:44

## R topics documented:

auta . . . . .	2
brca . . . . .	2
daneO . . . . .	3
daneSoc . . . . .	4
maratony . . . . .	4
mieszkania . . . . .	5
przezycia . . . . .	6
szkolne_wypadki . . . . .	7
titanic . . . . .	8
<b>Index</b>	<b>9</b>

---

auta

*Auta Dataset*

---

### **Description**

A subset of an auta2012 dataset from the package PogromcyDanych.

### **Usage**

```
data(auta)
```

### **Format**

a data.frame with 2400 rows and 8 columns

### **Details**

The subsequent columns describe:

- Marka Brand of the car
- Model Model of the car
- Cena Price
- KM Horse power
- Pojemnosc Volume of the engine
- Przebieg Mileage
- Paliwo Type of fuel
- Produkcja Year of production

### **Examples**

```
library("Przewodnik")  
summary(auta)
```

---

brca

*BRCA Dataset*

---

### **Description**

From The Cancer Genome Atlas dataset, subset for BRCA - BREast CANcer.

### **Usage**

```
data(brca)
```

**Format**

a data.frame with 695 rows and 16 columns

**Details**

The subsequent columns describe:

- time, death - survival status for patient
- subtype - cancer subtype
- p53mut - mutation in p53
- MDM2, ..., DNAJB12 - expression of different genes

**Examples**

```
library("Przewodnik")  
summary(brca)
```

---

daneO

*Crunching of Oncology Data*

---

**Description**

The subsequent columns describe:

- Wiek age
- Rozmiar.guza cancer size
- Wezly.chlonne lymph nodes
- Nowotwor cancer
- Receptory.estrogenowe estrogen receptors
- Receptory.progesteronowe progesteron receptors
- Niepowodzenia failures
- Okres.bez.wznowy time of observation
- VEGF vascular endothelial growth factor

**Usage**

```
data(daneO)
```

**Format**

a data.frame with 97 rows and 9 columns

**Examples**

```
library("Przewodnik")  
summary(daneO)
```

---

`daneSoc`*Example Sociodemographic Data*

---

**Description**

The subsequent columns describe:

- `wiek` Age
- `wyksztalcenie` Education
- `st.cywilny` Martial status
- `plec` Sex
- `praca` Work status
- `cisnienie.skurczowe` Blood pressure
- `cisnienie.rozkurczowe` Blood pressure

**Usage**

```
data(daneSoc)
```

**Format**

a data.frame with 204 rows and 7 columns

**Examples**

```
library("Przewodnik")  
summary(daneSoc)
```

---

`maratony`*Marathons in Warsaw / Poland*

---

**Description**

Data from Warsaw marathons 2013 - 2016.

**Usage**

```
data(maratony)
```

**Format**

a data.frame with 70736 rows and 14 columns

**Details**

Orlen Maraton Warszawski - Based on <https://www.orklenmarathon.pl/> database.

Polmaraton Warszawski - Based on <http://pzupolmaratonwarszawski.com/> database.

The subsequent columns describe:

- nazwisko\_imie First and Last name
- nr.startowy Number
- plec Gender
- rok Year
- biegi\_uliczne Name of the marathon
- czas\_brutto Gross Time
- czas\_brutto\_sec Gross Time in secs
- czas\_netto Net Time
- czas\_netto\_sec Net Time in sec
- grup\_wiek Age group
- kategoria Category
- kraj\_nazwa Country
- miejsce\_kat Place in Category
- miejsce\_open Place in Open

Prepared by Krzysztof Trajkowski.

**Examples**

```
library("Przewodnik")
summary(maratony)
```

---

mieszkania

*Housing Prices Data*

---

**Description**

The subsequent columns describe:

- cena price
- pokoi number of rooms
- powierzchnia surface
- dzielnica district
- typ.budynku house type

**Usage**

```
data(mieszkania)
```

**Format**

a data.frame with 200 rows and 5 columns

**Examples**

```
library("Przewodnik")
summary(mieszkania)
```

---

przezycia

*Mortality Data for Poland*

---

**Description**

Mortality data for Poland. Based on <http://www.mortality.org/> database. The subsequent columns describe:

- Year. Year or range of years (for both period and cohort data)
- Age. Age group for n-year interval from exact age x to just before exact age x+n, where n=1, 4, 5, or infinity (open age interval)
- m(x). Central death rate between ages x and x+n
- q(x). Probability of death between ages x and x+n
- a(x). Average length of survival between ages x and x+n for persons dying in the interval
- l(x). Number of survivors at exact age x, assuming  $l(0) = 100,000$
- d(x). Number of deaths between ages x and x+n
- L(x). Number of person-years lived between ages x and x+n
- T(x). Number of person-years remaining after exact age x
- e(x). Life expectancy at exact age x (in years)

**Usage**

```
data(przezycia)
```

**Format**

a data.frame with 11544 rows and 11 columns

**Details**

Find more at <http://www.mortality.org/Public/ExplanatoryNotes.php#CompleteDataSeries>

**Examples**

```
library("Przewodnik")
summary(przezycia)
```

---

szkolne_wypadki	<i>School Injuries in Poland</i>
-----------------	----------------------------------

---

## Description

Data about School Injuries in Poland in school years 2012/2013 and 2013/2014. Based on <http://www.cie.men.gov.pl/index.php/statystyczne/137.html>

## Usage

```
data(szkolne_wypadki)
```

## Format

a data.frame with 20958 rows and 7 columns

## Details

The subsequent columns describe:

- wojewodztwo Region
- szkola Type of school
- przypadek How heavy was the injury?
- rok.szkolny school year
- typ Is is about part of body or about type of injury
- rodzaj Detailed group
- liczba Number of injuries

Prepared by Krzysztof Trajkowski.

## Examples

```
library("Przewodnik")  
summary(szkolne_wypadki)
```

---

`titanic`*Titanic Dataset*

---

**Description**

Based on the titanic dataset from titanic package (from kaggle).

**Usage**

```
data(titanic)
```

**Format**

a data.frame with 891 rows and 6 columns

**Details**

The subsequent columns describe:

- Survived 1 - survived, 0 - death
- Pclass class, from 1 to 3
- Sex sex
- Age age
- Fare fare
- Embarked embarked

**Examples**

```
library("Przewodnik")  
summary(titanic)
```



# Index

## \*Topic **datasets**

auta, 2

brca, 2

dane0, 3

daneSoc, 4

maratony, 4

mieszkania, 5

przezycia, 6

szkolne\_wypadki, 7

titanic, 8

auta, 2

brca, 2

dane0, 3

daneSoc, 4

maratony, 4

mieszkania, 5

przezycia, 6

szkolne\_wypadki, 7

titanic, 8