# Package 'sampledatasets'

November 20, 2024

**Type** Package

**Title** A Collection of Sample Datasets

**Version** 0.1.0

**Maintainer** Renzo Caceres Rossi <arenzocaceresrossi@gmail.com>

**Description** Provides a collection of sample datasets
on various fields such as automotive performance and safety data to historical demographics and socioeconomic indicators, as well as recreational data.
It serves as a resource for researchers and analysts seeking to perform analyses and derive insights from classic data sets in R.

**License** GPL-3

**URL** https://github.com/lightbluetitan/sampledatasets

**Encoding** UTF-8

**LazyData** true

**Suggests** ggplot2, dplyr, knitr, rmarkdown, testthat (>= 3.0.0)

**Config/testthat/edition** 3

**VignetteBuilder** knitr

**RoxygenNote** 7.3.2

**NeedsCompilation** no

**Author** Renzo Caceres Rossi [aut, cre]

**Depends** R (>= 3.5.0)

**Repository** CRAN

**Date/Publication** 2024-11-20 19:00:02 UTC

# Contents

---

| arbuthnot_tbl | *Male and female births in London* |
|---|---|

---

#### Description

The 'arbuthnot_tbl' dataset contains historical baptism records collected by John Arbuthnot. It includes data on the number of boys and girls baptized each year from 1629 to 1710 in London. This dataset was historically significant in one of the earliest applications of statistical analysis to human populations. The dataset name has been suffixed with '_tbl' to indicate that it is a tibble, ensuring clarity within the 'sampledatasets' package.

#### Usage

```
data(arbuthnot_tbl)
```

#### Format

A tibble with 82 observations and 3 variables:

**year** Year of record (integer).

**boys** Number of boys baptized (integer).

**girls** Number of girls baptized (integer).

#### Source

Arbuthnot, J. (1710). *An argument for Divine Providence, taken from the constant regularity observed in the births of both sexes*.

---

| cards_tbl | *Standard Deck of Playing Cards* |
|---|---|

---

#### Description

The 'cards_tbl' dataset represents a standard deck of 52 playing cards. It includes information about the value, color, suit, and whether the card is a face card. The dataset name has been suffixed with '_tbl' to indicate that it is a tibble, ensuring clarity within the 'sampledatasets' package.

#### Usage

```
data(cards_tbl)
```

## Format

A tibble with 52 observations and 4 variables:

**value** Card value (factor with 13 levels: "2" to "10", "Jack", "Queen", "King", "Ace").

**color** Card color (factor with 2 levels: "black", "red").

**suit** Card suit (factor with 4 levels: "Clubs", "Diamonds", "Hearts", "Spades").

**face** Logical indicator (TRUE if the card is a face card, FALSE otherwise).

## Source

Constructed dataset representing a standard deck of cards.

---

cars_df                         *Speed and Stopping Distances of Cars*

---

## Description

The 'cars_df' dataset contains data on the speed of cars and the corresponding distances required to stop. This dataset is often used in statistical analysis and regression modeling. The dataset name has been suffixed with '_df' to indicate that it is a data frame, ensuring clarity within the 'sampledatasets' package.

## Usage

```
data(cars_df)
```

## Format

A data frame with 50 observations and 2 variables:

**speed** Speed of the car (mph) (numeric).

**dist** Stopping distance (ft) (numeric).

## Source

Data from the 1920s. Originally included in the *datasets* package in R.

---

| mtcars_df | *Motor Trend Car Road Tests* |
|---|---|

---

**Description**

The 'mtcars_df' dataset provides information extracted from the 1974 Motor Trend US magazine. It comprises data on fuel consumption and 10 aspects of automobile design and performance for 32 cars (1973–74 models). The dataset name has been suffixed with '_df' to identify it as a data frame, making it easier to distinguish within the 'sampledatasets' package.

**Usage**

```
data(mtcars_df)
```

**Format**

A data frame with 32 observations and 11 variables:

**mpg** Miles/(US) gallon (numeric).

**cyl** Number of cylinders (numeric).

**disp** Displacement (cu.in.) (numeric).

**hp** Gross horsepower (numeric).

**drat** Rear axle ratio (numeric).

**wt** Weight (1000 lbs) (numeric).

**qsec** 1/4 mile time (numeric).

**vs** Engine (0 = V-shaped, 1 = straight) (numeric).

**am** Transmission (0 = automatic, 1 = manual) (numeric).

**gear** Number of forward gears (numeric).

**carb** Number of carburetors (numeric).

**Source**

Motor Trend, 1974.

---

swiss_df *Swiss Fertility and Socioeconomic Indicators (1888)*

---

### Description

The 'swiss_df' dataset provides standardized fertility measures and socioeconomic indicators for 47 French-speaking provinces of Switzerland in 1888. The dataset name has been suffixed with '_df' to indicate that it is a data frame, ensuring clarity within the 'sampledatasets' package.

### Usage

```
data(swiss_df)
```

### Format

A data frame with 47 observations and 6 variables:

**Fertility**  Standardized fertility measure (numeric).

**Agriculture**  Percentage of males involved in agriculture as an occupation (numeric).

**Examination**  Percentage of draftees receiving the highest mark on an army examination (integer).

**Education**  Percentage of draftees with education beyond primary school (integer).

**Catholic**  Percentage of Catholic individuals (numeric).

**Infant.Mortality**  Infant mortality rate per 1000 live births (numeric).

### Source

Data from the Swiss Federal Statistical Office (1888).

---

usdatasets *sampledatasets: A Collection of Sample Datasets*

---

### Description

This package provides a variety of datasets related to automotive performance and safety data to historical demographics and socioeconomic indicators, as well as recreational data.

### Details

sampledatasets: A Collection of Sample Datasets

A Collection of Sample Datasets.

### Author(s)

**Maintainer**: Renzo Cáceres Rossi <arenzocaceresrossi@gmail.com>

**See Also**

Useful links:

- [https://github.com/lightbluetitan/LatinR_2024_R_DEVELOPER](https://github.com/lightbluetitan/LatinR_2024_R_DEVELOPER)

# Index