

# Package ‘CardioDataSets’

May 13, 2025

**Type** Package

**Title** A Comprehensive Collection of Cardiovascular and Heart Disease Datasets

**Version** 0.1.0

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

**Description** Offers a diverse collection of datasets focused on cardiovascular and heart disease research, including heart failure, myocardial infarction, aortic dissection, transplant outcomes, cardiovascular risk factors, drug efficacy, and mortality trends.

Designed for researchers, clinicians, epidemiologists, and data scientists, the package features clinical, epidemiological, and simulated datasets covering a wide range of conditions and treatments such as statins, anticoagulants, and beta blockers.

It supports analyses related to disease progression, treatment effects, rehospitalization, and public health outcomes across various cardiovascular patient populations.

**License** GPL-3

**URL** <https://github.com/lightbluetitan/cardiодatasets>,  
<https://lightbluetitan.github.io/cardiодatasets/>

**BugReports** <https://github.com/lightbluetitan/cardiодatasets/issues>

**Encoding** UTF-8

**LazyData** true

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

**Depends** R (>= 4.2.0)

**Imports** utils

**RoxygenNote** 7.3.2

**Config/testthat/edition** 3

**VignetteBuilder** knitr

**NeedsCompilation** no

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

**Repository** CRAN

**Date/Publication** 2025-05-13 08:20:06 UTC

## Contents

acs_patients_df . . . . .	2
age_heartrate_df . . . . .	4
ami_occurrences_tbl_df . . . . .	4
aortaDiss_tbl_df . . . . .	5
betablockers_matrix . . . . .	6
cad_anticoagulants_df . . . . .	7
cardiac_failure_df . . . . .	8
cardiac_gwas_df . . . . .	9
CardioDataSets . . . . .	10
cardioRiskFactors_df . . . . .	10
cardiovascular_list . . . . .	11
cardio_diabetes_tbl_df . . . . .	12
coronary_death_df . . . . .	13
cpr_survival_tbl_df . . . . .	14
cv_mortality_ts . . . . .	14
emotion_heartrate_df . . . . .	15
heartdiseaserisk_tbl_df . . . . .	16
heartdisease_tbl_df . . . . .	17
heartfailure_df . . . . .	18
heartTransplantTime_tbl_df . . . . .	19
heart_transplant_df . . . . .	19
hfPrevention_mtc_network . . . . .	20
mriCardioVars_tbl_df . . . . .	21
muscatine_coronary_risk_df . . . . .	23
myocardialinfarction_df . . . . .	24
patient_CAV_df . . . . .	25
radial_ivus_df . . . . .	25
scottish_CVD_df . . . . .	27
statinMIrisk_df . . . . .	28
sulphinpyrazone_tbl_df . . . . .	28
usMortality_df . . . . .	29
view_datasets . . . . .	30
<b>Index</b>	<b>31</b>

---

acs_patients_df	<i>Acute Coronary Syndrome (ACS) Patient Data</i>
-----------------	---

---

### Description

This dataset, `acs_patients_df`, is a data frame containing demographic and clinical data from 857 patients with Acute Coronary Syndrome (ACS). It includes 17 variables covering patient characteristics, vital signs, laboratory results, and risk factors.

**Usage**

```
data(acs_patients_df)
```

**Format**

A data frame with 857 observations and 17 variables:

**age** Patient age in years (integer)

**sex** Patient sex (character)

**cardiogenicShock** Presence of cardiogenic shock (character)

**entry** Method of hospital entry (character)

**Dx** Diagnosis (character)

**EF** Ejection fraction percentage (numeric)

**height** Height in cm (numeric)

**weight** Weight in kg (numeric)

**BMI** Body Mass Index in kg/m<sup>2</sup> (numeric)

**obesity** Obesity status (character)

**TC** Total cholesterol in mg/dL (numeric)

**LDLC** LDL cholesterol in mg/dL (integer)

**HDLC** HDL cholesterol in mg/dL (integer)

**TG** Triglycerides in mg/dL (integer)

**DM** Diabetes mellitus status (character)

**HBP** High blood pressure status (character)

**smoking** Smoking status (character)

**Details**

The dataset name has been kept as 'acs\_patients\_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the CardioDataSets package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a standard data frame. The original content has not been modified in any way.

**Source**

Data taken from the moonBook package version 0.3.1

---

age\_hearttrate\_df      *Age vs. Maximum Heart Rate*

---

### Description

This dataset, `age_hearttrate_df`, is a data frame containing simulated data representing the relationship between age and maximum heart rate. It includes 15 observations based on established physiological models.

### Usage

```
data(age_hearttrate_df)
```

### Format

A data frame with 15 observations and 2 variables:

**age** Age in years (numeric)

**maxrate** Maximum predicted heart rate in beats per minute (numeric)

### Details

The dataset name has been kept as `'age_hearttrate_df'` to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the `CardioDataSets` package and assists users in identifying its specific characteristics. The suffix `'df'` indicates that the dataset is a standard data frame. The original content has not been modified in any way.

### Source

Data taken from the `UsingR` package version 2.0-7. Original research: Tanaka H, Monahan KD, Seals DR (2001). "Age-predicted maximal heart rate revisited." *Journal of the American College of Cardiology*, 37(1):153-156.

---

ami\_occurrences\_tbl\_df      *Acute Myocardial Infarction (Heart Attack) Events*

---

### Description

This dataset, `ami_occurrences_tbl_df`, is a tibble containing simulated but realistic daily counts of Acute Myocardial Infarction (AMI) occurrences in New York City over one year (365 days). The data represents the number of heart attack events recorded each day.

### Usage

```
data(ami_occurrences_tbl_df)
```

**Format**

A tibble with 365 observations and 1 variable:

**ami** Number of Acute Myocardial Infarction events recorded each day (integer vector)

**Details**

The dataset name has been kept as 'ami\_occurrences\_tbl\_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the CardioDataSets package and assists users in identifying its specific characteristics. The suffix 'tbl\_df' indicates that the dataset is a tibble. The original content has not been modified in any way.

**Source**

Data taken from the openintro package version 2.5.0

---

aortaDiss_tbl_df	<i>Aortic dissection patients</i>
------------------	-----------------------------------

---

**Description**

This dataset, aortaDiss\_tbl\_df, is a tibble containing clinical information from 226 patients with aortic dissection. It includes demographic variables, symptom presentation, and risk factor data.

**Usage**

```
data(aortaDiss_tbl_df)
```

**Format**

A tibble with 226 observations and 10 variables:

**Gender** Patient gender (numeric)

**Age** Patient age in years (numeric)

**Age\_C** Categorized age (numeric)

**Aortadis** Aortic dissection status (numeric)

**Acute** Acute presentation indicator (numeric)

**Acute3** Three-level acute presentation classification (numeric)

**Stomach\_Ache** Presence of stomach ache (numeric)

**Hyper** Hypertension status (numeric)

**Smoking** Smoking status (numeric)

**Radiation** Radiation exposure (numeric)

## Details

The dataset name has been kept as 'aortaDiss\_tbl\_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the CardioDataSets package and assists users in identifying its specific characteristics. The suffix 'tbl\_df' indicates that the dataset is a tibble. The original content has not been modified in any way.

## Source

Data taken from the psfmi package version 1.4.0

---

betablockers\_matrix     *FDA Beta Blockers Adverse Events*

---

## Description

This dataset, betablockers\_matrix, is a matrix containing adverse event reports from the FDA Adverse Event Reporting System (FAERS) for 9 beta blockers from Q1 2021 to Q4 2023. The matrix includes 501 adverse events (rows) across 9 medications (columns).

## Usage

```
data(betablockers_matrix)
```

## Format

A matrix with 501 rows (adverse events) and 9 columns (beta blockers):

**Acebutolol** Adverse event counts for Acebutolol (integer)

**Atenolol** Adverse event counts for Atenolol (integer)

**Bisoprolol** Adverse event counts for Bisoprolol (integer)

**Carvedilol** Adverse event counts for Carvedilol (integer)

**Metoprolol** Adverse event counts for Metoprolol (integer)

**Nadolol** Adverse event counts for Nadolol (integer)

**Propranolol** Adverse event counts for Propranolol (integer)

**Timolol** Adverse event counts for Timolol (integer)

**Other** Adverse event counts for other beta blockers (integer)

## Details

The dataset name has been kept as 'betablockers\_matrix' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the CardioDataSets package and assists users in identifying its specific characteristics. The suffix 'matrix' indicates that the dataset is a matrix object. The original content has not been modified in any way.

**Source**

Data taken from the MDDC package version 1.1.0. Original data: FDA Adverse Event Reporting System (FAERS) database, Q1 2021 to Q4 2023.

---

cad\_anticoagulants\_df *Anticoagulants for CAD Patients*

---

**Description**

This dataset, cad\_anticoagulants\_df, is a data frame containing information from 34 clinical trials examining the effectiveness of oral anticoagulants in patients with coronary artery disease. It includes data on treatment outcomes comparing anticoagulant therapy with control groups.

**Usage**

```
data(cad_anticoagulants_df)
```

**Format**

A data frame with 34 observations and 9 variables:

**study** Study identifier (character vector)

**year** Year of publication (integer vector)

**intensity** Intensity of anticoagulation treatment (character vector)

**asp.t** Aspirin use in treatment group (integer vector)

**asp.c** Aspirin use in control group (integer vector)

**ai** Number of events in treatment group (integer vector)

**n1i** Total number of participants in treatment group (integer vector)

**ci** Number of events in control group (integer vector)

**n2i** Total number of participants in control group (integer vector)

**Details**

The dataset name has been kept as 'cad\_anticoagulants\_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the CardioDataSets package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a standard data frame. The original content has not been modified in any way.

**Source**

Data taken from the metadat package version 1.2-0

---

cardiac\_failure\_df      *Heart Failure Clinical Dataset*

---

## Description

This dataset, `cardiac_failure_df`, is a data frame containing clinical data from 299 patients with heart failure. It includes 13 variables covering demographic information, medical history, laboratory results, and mortality outcomes.

## Usage

```
data(cardiac_failure_df)
```

## Format

A data frame with 299 observations and 13 variables:

**age** Patient age in years (numeric)  
**anaemia** Presence of anaemia (integer: 0=no, 1=yes)  
**creatinine\_phosphokinase** Level of CPK enzyme in mcg/L (integer)  
**diabetes** Presence of diabetes (integer: 0=no, 1=yes)  
**ejection\_fraction** Percentage of blood leaving heart (integer)  
**high\_blood\_pressure** Presence of hypertension (integer: 0=no, 1=yes)  
**platelets** Platelet count in kiloplatelets/mL (numeric)  
**serum\_creatinine** Level of serum creatinine in mg/dL (numeric)  
**serum\_sodium** Level of serum sodium in mEq/L (integer)  
**sex** Patient sex (integer: 0=female, 1=male)  
**smoking** Smoking status (integer: 0=no, 1=yes)  
**time** Follow-up period in days (integer)  
**DEATH\_EVENT** Death during follow-up (integer: 0=no, 1=yes)

## Details

The dataset name has been kept as `'cardiac_failure_df'` to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the `CardioDataSets` package and assists users in identifying its specific characteristics. The suffix `'df'` indicates that the dataset is a standard data frame. The original content has not been modified in any way.

## Source

Data taken from the `SOPC` package version 0.1.0



---

cardiac_gwas_df	<i>Coronary Artery Disease GWAS Meta-Analysis</i>
-----------------	---

---

## Description

This dataset, `cardiac_gwas_df`, is a data frame containing genome-wide association study (GWAS) results from a multi-ethnic meta-analysis of coronary artery disease (CAD). It includes 9,919 genetic variants with their effect sizes and study characteristics.

## Usage

```
data(cardiac_gwas_df)
```

## Format

A data frame with 9,919 observations and 7 variables:

**beta\_flipped** Effect size estimates (numeric)  
**gcse** Genomic control standard error (numeric)  
**variants** Genetic variant identifiers (character)  
**studies** Participating studies (character)  
**cases** Number of cases (integer)  
**controls** Number of controls (integer)  
**fdr214\_gwas46** False discovery rate adjusted p-values (numeric)

## Details

The dataset name has been kept as `'cardiac_gwas_df'` to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the `CardioDataSets` package and assists users in identifying its specific characteristics. The suffix `'df'` indicates that the dataset is a standard data frame. The original content has not been modified in any way.

## Source

Data taken from the `getmstatistic` package version 0.2.2

---

CardioDataSets      *CardioDataSets: A Comprehensive Collection of Cardiovascular and Heart Disease Datasets*

---

### Description

This package provides a wide variety of datasets focused on heart and cardiovascular research, covering heart disease, myocardial infarction, heart failure, stroke, ischemic heart disease, risk factors, clinical trials, and treatment outcomes.

### Details

CardioDataSets: A Comprehensive Collection of Cardiovascular and Heart Disease Datasets  
A Comprehensive Collection of Cardiovascular and Heart Disease Datasets.

### Author(s)

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

### See Also

Useful links:

- <https://github.com/lightbluetitan/cardiодatasets>

---

cardioRiskFactors\_df      *Cardiovascular Risk Factors*

---

### Description

This dataset, cardioRiskFactors\_df, is a data frame containing information from a study investigating the association between uric acid and cardiovascular risk factors in developing countries. It includes data from 998 participants (474 men and 524 women) aged 25-64 years.

### Usage

```
data(cardioRiskFactors_df)
```

### Format

A data frame with 998 observations and 14 variables:

**age** Age in years (integer)

**bmi** Body Mass Index in kg/m<sup>2</sup> (numeric)

**waisthip** Waist-to-hip ratio (numeric)

**smok** Smoking status (integer)  
**choles** Total cholesterol in mg/dL (numeric)  
**trig** Triglycerides in mg/dL (numeric)  
**hdl** HDL cholesterol in mg/dL (numeric)  
**ldl** LDL cholesterol in mg/dL (numeric)  
**sys** Systolic blood pressure in mmHg (integer)  
**dia** Diastolic blood pressure in mmHg (numeric)  
**Uric** Uric acid level in mg/dL (integer)  
**sex** Sex (integer)  
**alco** Alcohol consumption (numeric)  
**apoa** Apolipoprotein A in mg/dL (numeric)

### Details

The dataset name has been kept as 'cardioRiskFactors\_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the CardioDataSets package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a standard data frame. The original content has not been modified in any way.

### Source

Data taken from the Rfit package version 0.27.0. Original study: Heritier S, Cantoni E, Copt S, Victoria-Feser M (2009). Robust Methods in Biostatistics. New York: John Wiley and Sons.

---

cardiovascular\_list     *Statin Dose Comparison Trials for CVD*

---

### Description

This dataset, cardiovascular\_list, is a list containing data from 34 clinical trials comparing low dose (1), high dose (2), and placebo (3) statins for cardiovascular disease prevention. The dataset includes study identifiers, treatment assignments, and outcome counts.

### Usage

```
data(cardiovascular_list)
```

### Format

A list with 4 components:

**Study** Study identifiers (integer vector of length 34)

**Treat** Treatment assignments (numeric vector: 1=low dose, 2=high dose, 3=placebo)

**Outcomes** Outcome matrix with 34 rows and 3 columns:

**Alive** Number of patients alive (numeric)

**FnCVD** Number with non-fatal CVD events (numeric)

**FCVD** Number with fatal CVD events (numeric)

**N** Sample sizes (numeric vector of length 34)

### Details

The dataset name has been kept as 'cardiovascular\_list' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the CardioDataSets package and assists users in identifying its specific characteristics. The original content has not been modified in any way.

### Source

Data taken from the bnma package version 1.6.0

---

cardio\_diabetes\_tbl\_df

*Cardiovascular risks of diabetes drugs*

---

### Description

This dataset, cardio\_diabetes\_tbl\_df, is a tibble containing information comparing cardiovascular problems between two diabetes medications (Rosiglitazone and Pioglitazone) in elderly Medicare patients. It includes data from 227,571 patients.

### Usage

```
data(cardio_diabetes_tbl_df)
```

### Format

A tibble with 227,571 observations and 2 variables:

**treatment** Type of diabetes medication (factor with 2 levels: Rosiglitazone or Pioglitazone)

**cardiovascular\_problems** Presence of cardiovascular problems (factor with 2 levels)

### Details

The dataset name has been kept as 'cardio\_diabetes\_tbl\_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the CardioDataSets package and assists users in identifying its specific characteristics. The suffix 'tbl\_df' indicates that the dataset is a tibble. The original content has not been modified in any way.

### Source

Data taken from the openintro package version 2.5.0. Original study: Graham DJ, et al. (2010). "Risk of acute myocardial infarction, stroke, heart failure, and death in elderly Medicare patients treated with rosiglitazone or pioglitazone." JAMA, 304(4):411.

---

coronary\_death\_df      *High vs Moderate Statins for MI Prevention*

---

### Description

This dataset, `coronary_death_df`, is a data frame containing information from 4 clinical trials comparing intensive (high dose) versus moderate (standard dose) statin therapy for preventing coronary death or myocardial infarction. It includes data on treatment outcomes across multiple endpoints.

### Usage

```
data(coronary_death_df)
```

### Format

A data frame with 4 observations and 16 variables:

**trial** Trial identifier (character vector)  
**pop** Patient population description (character vector)  
**nt** Number of patients in treatment group (integer vector)  
**nc** Number of patients in control group (integer vector)  
**ep1t** Endpoint 1 events in treatment group (integer vector)  
**ep1c** Endpoint 1 events in control group (integer vector)  
**ep2t** Endpoint 2 events in treatment group (integer vector)  
**ep2c** Endpoint 2 events in control group (integer vector)  
**ep3t** Endpoint 3 events in treatment group (integer vector)  
**ep3c** Endpoint 3 events in control group (integer vector)  
**ep4t** Endpoint 4 events in treatment group (integer vector)  
**ep4c** Endpoint 4 events in control group (integer vector)  
**ep5t** Endpoint 5 events in treatment group (integer vector)  
**ep5c** Endpoint 5 events in control group (integer vector)  
**ep6t** Endpoint 6 events in treatment group (integer vector)  
**ep6c** Endpoint 6 events in control group (integer vector)

### Details

The dataset name has been kept as `'coronary_death_df'` to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the `CardioDataSets` package and assists users in identifying its specific characteristics. The suffix `'df'` indicates that the dataset is a standard data frame. The original content has not been modified in any way.

### Source

Data taken from the `metadat` package version 1.2-0

---

cpr\_survival\_tbl\_df     *Blood thinners in CPR survival*

---

### Description

This dataset, `cpr_survival_tbl_df`, is a tibble containing information from a study examining the effect of blood thinners on survival rates in CPR patients. The study randomly assigned 90 patients to either receive a blood thinner (treatment group) or not receive one (control group), with the outcome being survival for at least 24 hours.

### Usage

```
data(cpr_survival_tbl_df)
```

### Format

A tibble with 90 observations and 2 variables:

**group** Treatment assignment (factor with 2 levels: "control" and "treatment")

**outcome** Survival status (factor with 2 levels: "died" and "survived")

### Details

The dataset name has been kept as `'cpr_survival_tbl_df'` to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the CardioDataSets package and assists users in identifying its specific characteristics. The suffix `'tbl_df'` indicates that the dataset is a tibble. The original content has not been modified in any way.

### Source

Data taken from the `openintro` package version 2.5.0

---

cv\_mortality\_ts     *LA pollution and cardiovascular mortality*

---

### Description

This dataset, `cv_mortality_ts`, is a time series containing weekly cardiovascular mortality data from Los Angeles County. It consists of 508 six-day smoothed averages obtained by filtering daily values over the 10-year period from 1970 to 1979.

### Usage

```
data(cv_mortality_ts)
```

**Format**

A time series object (ts) with 508 observations:

**cv\_mortality** Weekly cardiovascular mortality counts (numeric vector)

**Details**

The dataset name has been kept as 'cv\_mortality\_ts' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the CardioDataSets package and assists users in identifying its specific characteristics. The suffix 'ts' indicates that the dataset is a time series object. The original content has not been modified in any way.

Time series characteristics: - Start: 1970, Week 1 - End: 1979, Week 40 - Frequency: 52 (weekly data)

**Source**

Data taken from the astsa package version 2.2

---

emotion\_hearttrate\_df *Anger recall effect on heart rate (Lakens, 2013)*

---

**Description**

This dataset, emotion\_hearttrate\_df, is a data frame containing heart rate measurements from a study investigating how recalling anger affects heart rate. It includes baseline and anger-induced heart rate measurements from 68 participants.

**Usage**

```
data(emotion_hearttrate_df)
```

**Format**

A data frame with 68 observations and 3 variables:

**ID** Participant identification number (integer vector)

**HR\_baseline** Baseline heart rate in beats per minute (numeric vector)

**HR\_anger** Heart rate during anger recall in beats per minute (numeric vector)

**Details**

The dataset name has been kept as 'emotion\_hearttrate\_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the CardioDataSets package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a standard data frame. The original content has not been modified in any way.

**Source**

Data taken from the `esci` package version 1.0-7. Original study: Lakens D (2013). Conceptual replication of Ekman et al. (1983) emotion study.

---

heartdiseaserisk\_tbl\_df

*Heart Disease Risk Factors*

---

**Description**

This dataset, `heartdiseaserisk_tbl_df`, is a tibble containing cardiovascular risk factor data from 498 individuals. It includes measures of physical activity (biking), smoking habits, and heart disease prevalence.

**Usage**

```
data(heartdiseaserisk_tbl_df)
```

**Format**

A tibble with 498 observations and 3 variables:

**Biking** Frequency of biking activity (numeric)

**Heart.disease** Prevalence of heart disease (numeric)

**Smoking** Smoking frequency or intensity (numeric)

**Details**

The dataset name has been kept as `'heartdiseaserisk_tbl_df'` to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the `CardioDataSets` package and assists users in identifying its specific characteristics. The suffix `'tbl_df'` indicates that the dataset is a tibble. The original content has not been modified in any way.

**Source**

Data taken from the `Path.Analysis` package version 0.1



---

heartdisease\_tbl\_df     *Heart Disease Patients Clinical Data*

---

### Description

This dataset, `heartdisease_tbl_df`, is a tibble containing information on individuals evaluated for heart disease. It is a cleaned version of the original "Heart Disease" dataset from the UCI Machine Learning Repository, and includes 303 observations on 9 variables.

### Usage

```
data(heartdisease_tbl_df)
```

### Format

A tibble with 303 observations and 9 variables:

**Age** Age of the individual (numeric).

**Sex** Sex of the individual (factor with 2 levels: typically "Male" and "Female").

**ChestPain** Type of chest pain experienced (factor with 4 levels).

**BP** Resting blood pressure (numeric).

**Cholesterol** Serum cholesterol in mg/dl (numeric).

**BloodSugar** Indicates if fasting blood sugar > 120 mg/dl (logical).

**MaximumHR** Maximum heart rate achieved (numeric).

**ExerciseInducedAngina** Exercise-induced angina (factor with 2 levels).

**HeartDisease** Presence or absence of heart disease (factor with 2 levels).

### Details

The dataset name has been kept as `'heartdisease_tbl_df'` to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the `CardioDataSets` package and assists users in identifying its specific characteristics. The suffix `'tbl_df'` indicates that the dataset is a tibble. The original content has not been modified in any way.

### Source

Data taken from the `cheese` package version 0.1.2. Original source: UCI Machine Learning Repository. Heart Disease Data Set. <https://archive.ics.uci.edu/ml/datasets/Heart+Disease>

---

heartfailure_df	<i>Heart Failure rehospitalization risk</i>
-----------------	---

---

## Description

This dataset, `heartfailure_df`, is a data frame containing simulated data from 800 patients with heart failure who are at risk of recurrent hospitalization. The dataset includes 3,068 observations (2,268 events) tracking patient outcomes over time.

## Usage

```
data(heartfailure_df)
```

## Format

A data frame with 3,068 observations and 6 variables:

- id** Patient identification number (integer vector)
- treatment** Treatment assignment (factor with 2 levels)
- t0** Start time of observation period (numeric vector)
- t1** End time of observation period (numeric vector)
- enum** Event number (numeric vector)
- event** Event indicator (numeric vector)

## Details

The dataset name has been kept as `'heartfailure_df'` to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the `CardioDataSets` package and assists users in identifying its specific characteristics. The suffix `'df'` indicates that the dataset is a standard data frame. The original content has not been modified in any way.

## Source

Data taken from the `survPen` package version 2.0-2. Based on `hfaction_cpx12` dataset from package `WA`.

---

heartTransplantTime\_tbl\_df

*Artificial Heart Transplant Durations*

---

### Description

This dataset, heartTransplantTime\_tbl\_df, is a tibble containing the durations (in hours) of 15 artificial heart transplant operations.

### Usage

```
data(heartTransplantTime_tbl_df)
```

### Format

A tibble with 15 observations and 1 variable:

**duration** Operation duration in hours (numeric)

### Details

The dataset name has been kept as 'heartTransplantTime\_tbl\_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the CardioDataSets package and assists users in identifying its specific characteristics. The suffix 'tbl\_df' indicates that the dataset is a tibble. The original content has not been modified in any way.

### Source

Data taken from the BSDA package version 1.2.3. Original source: Kitchens LJ (2003). "Basic Statistics and Data Analysis." Pacific Grove, CA: Brooks/Cole, a division of Thomson Learning.

---

heart\_transplant\_df

*Stanford Heart Transplant Data*

---

### Description

This dataset, heart\_transplant\_df, is a data frame containing survival data from the Stanford heart transplant program. It includes information on 172 patients with follow-up times, transplant status, and clinical covariates.

### Usage

```
data(heart_transplant_df)
```

**Format**

A data frame with 172 observations and 8 variables:

- start** Start time of interval (numeric)
- stop** End time of interval (numeric)
- event** Survival status (numeric: 1=event, 0=censored)
- age** Patient age at enrollment (numeric)
- year** Year of enrollment (numeric)
- surgery** Prior bypass surgery (numeric)
- transplant** Transplant status (factor: 0=no, 1=yes)
- id** Patient identification number (numeric)

**Details**

The dataset name has been kept as 'heart\_transplant\_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the CardioDataSets package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a standard data frame. The original content has not been modified in any way.

**Source**

Data taken from the lrsat package version 0.2.13. Original source: Stanford Heart Transplant Study data from the survival package.

---

hfPrevention\_mtc\_network

*Statins for Heart Failure Prevention*

---

**Description**

This dataset, hfPrevention\_mtc\_network, contains network meta-analysis data from 19 trials comparing statins versus placebo or usual care for cholesterol lowering in heart failure. The main outcome measured is the number of deaths. Trials are categorized as either primary prevention (no previous heart disease) or secondary prevention (previous heart disease).

**Usage**

```
data(hfPrevention_mtc_network)
```

**Format**

An 'mtc.network' object (list) with 4 components:

**description** Character string describing the analysis: "Cholesterol lowering in HF (outcome: death)"

**treatments** Data frame with 2 treatments:

**id** Treatment ID (factor with 2 levels)

**description** Treatment description (character vector)

**data.ab** Data frame with 38 rows (arm-level data):

**study** Study ID (factor with 19 levels)

**treatment** Treatment assignment (factor with 2 levels)

**responders** Number of deaths (integer vector)

**sampleSize** Total sample size per arm (integer vector)

**studies** Data frame with 19 rows (study-level data):

**study** Study ID (factor with 19 levels)

**secondary** Prevention type: 0 = primary, 1 = secondary (integer vector)

**Details**

The dataset name has been kept as 'hfPrevention\_mtc\_network' to maintain consistency with its original source and to avoid confusion with other datasets. This naming convention helps identify this specific network meta-analysis dataset from the CardioDataSets package. The dataset is structured as an 'mtc.network' object, which is the standard format for network meta-analysis in the gemtc package. The original content has not been modified.

**Source**

Data taken from the gemtc package version 1.0-2. Original publication: Dias S, Sutton AJ, Welton NJ, Ades AE (2013). "Heterogeneity - Subgroups, Meta-Regression, Bias, and Bias-Adjustment." Medical Decision Making, 33(5):618-640.

---

mriCardioVars\_tbl\_df *Elderly CV/MRI and Biomarkers*

---

**Description**

This dataset, mriCardioVars\_tbl\_df, is a tibble containing MRI and clinical data from 735 elderly participants in a U.S. observational study of cardiovascular and cerebrovascular disease incidence. It includes 30 variables covering demographic, clinical, and imaging measures.

**Usage**

```
data(mriCardioVars_tbl_df)
```

**Format**

A tibble with 735 observations and 30 variables:

**ptid** Patient identification number (numeric)  
**mridate** MRI date (Date)  
**age** Age in years (numeric)  
**sex** Sex (character)  
**race** Race (character)  
**weight** Weight in kg (numeric)  
**height** Height in cm (numeric)  
**packyrs** Smoking pack-years (numeric)  
**yrsquit** Years since quitting smoking (numeric)  
**alcoh** Alcohol consumption (numeric)  
**physact** Physical activity level (numeric)  
**chf** Congestive heart failure status (numeric)  
**chd** Coronary heart disease status (numeric)  
**stroke** Stroke history (numeric)  
**diabetes** Diabetes status (numeric)  
**genhlth** General health status (numeric)  
**ldl** LDL cholesterol in mg/dL (numeric)  
**alb** Albumin level (numeric)  
**crt** Creatinine level (numeric)  
**plt** Platelet count (numeric)  
**sbp** Systolic blood pressure in mmHg (numeric)  
**aai** Ankle-arm index (numeric)  
**fev** Forced expiratory volume (numeric)  
**dsst** Digit Symbol Substitution Test score (numeric)  
**atrophy** Brain atrophy measure (numeric)  
**whgrd** White matter hyperintensity grade (numeric)  
**numinf** Number of brain infarcts (numeric)  
**volinf** Volume of brain infarcts (numeric)  
**obstime** Observation time (numeric)  
**death** Mortality status (numeric)

**Details**

The dataset name has been kept as 'mriCardioVars\_tbl\_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the Cardio-DataSets package and assists users in identifying its specific characteristics. The suffix 'tbl\_df' indicates that the dataset is a tibble. The original content has not been modified in any way.

**Source**

Data taken from the `rigr` package version 1.0.7

---

muscatine\_coronary\_risk\_df  
*Muscatine pediatric CRF*

---

**Description**

This dataset, `muscatine_coronary_risk_df`, is a data frame containing longitudinal observations from the Muscatine Coronary Risk Factor (MCRF) study, which examined the development of coronary disease risk factors in children. It includes 14,568 observations of 4,856 children tracked from 1977 to 1981.

**Usage**

```
data(muscatine_coronary_risk_df)
```

**Format**

A data frame with 14,568 observations and 7 variables:

**id** Child identification number (integer)  
**gender** Gender of child (factor with 2 levels)  
**base\_age** Age at first observation in years (integer)  
**age** Current age in years (integer)  
**occasion** Measurement occasion (integer)  
**obese** Obesity status (factor with 2 levels)  
**numobese** Numeric obesity indicator (numeric)

**Details**

The dataset name has been kept as `'muscatine_coronary_risk_df'` to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the `CardioDataSets` package and assists users in identifying its specific characteristics. The suffix `'df'` indicates that the dataset is a standard data frame. The original content has not been modified in any way.

**Source**

Data taken from the `geepack` package version 1.3.12. Original study: The Muscatine Coronary Risk Factor Study, University of Iowa, 1977-1981.

---

myocardialinfarction\_df

*Streptokinase Therapy in AMI*

---

## Description

This dataset, `myocardialinfarction_df`, is a data frame containing information from 33 clinical trials comparing intravenous streptokinase versus placebo or no therapy in patients hospitalized for acute myocardial infarction. It includes data on treatment outcomes between intervention and control groups.

## Usage

```
data(myocardialinfarction_df)
```

## Format

A data frame with 33 observations and 6 variables:

**trial** Trial identifier (character vector)

**year** Year of publication (integer vector)

**ai** Number of events in treatment group (integer vector)

**n1i** Total number of participants in treatment group (integer vector)

**ci** Number of events in control group (integer vector)

**n2i** Total number of participants in control group (integer vector)

## Details

The dataset name has been kept as `'myocardialinfarction_df'` to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the `CardioDataSets` package and assists users in identifying its specific characteristics. The suffix `'df'` indicates that the dataset is a standard data frame. The original content has not been modified in any way.

## Source

Data taken from the `metadat` package version 1.2-0. Original publication: Lau J, Antman EM, Jimenez-Silva J, Kupelnick B, Mosteller F, Chalmers TC (1992). "Cumulative meta-analysis of therapeutic trials for myocardial infarction." *New England Journal of Medicine*, 327(4):248-254.



---

patient_CAV_df	<i>CAV in Heart Transplant Patients</i>
----------------	---

---

**Description**

This dataset, `patient_CAV_df`, is a data frame containing longitudinal follow-up data from heart transplant recipients at Papworth Hospital, UK. It tracks 2,803 angiographic examinations for the onset of cardiac allograft vasculopathy and mortality.

**Usage**

```
data(patient_CAV_df)
```

**Format**

A data frame with 2,803 observations and 5 variables:

**PTNUM** Patient identification number (integer)

**years** Time since transplant in years (numeric)

**state** Disease state (numeric)

**dage** Donor age in years (integer)

**pdiag** Primary diagnosis code (numeric)

**Details**

The dataset name has been kept as 'patient\_CAV\_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the `CardioDataSets` package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a standard data frame. The original content has not been modified in any way.

**Source**

Data taken from the `flexmsm` package version 0.1.2. Original data: Papworth Hospital, UK. Subset of cav data from `msm` package.

---

radial_ivus_df	<i>Radial Artery IVUS Patient Data</i>
----------------	--

---

**Description**

This dataset, `radial_ivus_df`, is a data frame containing demographic and clinical data from 115 patients who underwent intravascular ultrasound (IVUS) examination of the radial artery following transradial coronary angiography. It includes 15 variables covering patient characteristics, laboratory results, and IVUS measurements.

**Usage**

```
data(radial_ivus_df)
```

**Format**

A data frame with 115 observations and 15 variables:

**male** Male sex indicator (integer: 0/1)

**age** Age in years (integer)

**height** Height in cm (numeric)

**weight** Weight in kg (numeric)

**HBP** High blood pressure status (integer: 0/1)

**DM** Diabetes mellitus status (integer: 0/1)

**smoking** Smoking status (factor with 3 levels)

**TC** Total cholesterol in mg/dL (integer)

**TG** Triglycerides in mg/dL (integer)

**HDL** HDL cholesterol in mg/dL (integer)

**LDL** LDL cholesterol in mg/dL (integer)

**hsCRP** High-sensitivity C-reactive protein in mg/L (numeric)

**NTAV** Normalized total atheroma volume (numeric)

**PAV** Percent atheroma volume (numeric)

**sex** Sex (factor with 2 levels)

**Details**

The dataset name has been kept as 'radial\_ivus\_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the CardioDataSets package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a standard data frame. The original content has not been modified in any way.

**Source**

Data taken from the moonBook package version 0.3.1

---

scottish_CVD_df	<i>Scottish Health Survey CVD</i>
-----------------	-----------------------------------

---

## Description

This dataset, `scottish_CVD_df`, is a data frame containing cardiovascular health data from the 1998 Scottish Health Survey. It includes information from 8,804 respondents aged 18-64, with variables covering demographics, health behaviors, and cardiovascular disease status.

## Usage

```
data(scottish_CVD_df)
```

## Format

A data frame with 8,804 observations and 8 variables:

- age** Respondent age in years (integer)
- sex** Respondent sex (factor with 2 levels)
- sc** Social class (factor with 3 levels)
- cvddef** Doctor-diagnosed CVD status (integer: 0=no, 1=yes)
- carstair** Carstairs deprivation score (numeric)
- smoke** Smoking status (factor with 5 levels)
- id** Respondent identification number (integer)
- area** Geographic area code (integer)

## Details

The dataset name has been kept as `'scottish_CVD_df'` to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the `CardioDataSets` package and assists users in identifying its specific characteristics. The suffix `'df'` indicates that the dataset is a standard data frame. The original content has not been modified in any way.

## Source

Data taken from the `R2MLwiN` package version 0.8-9. Original survey: 1998 Scottish Health Survey. Methodology reference: Charlton C, Rasbash J, Browne WJ, Healy M, Cameron B (2024). `MLwiN` Version 3.09. Centre for Multilevel Modelling, University of Bristol.

---

statinMIrisk_df	<i>Statin intensity and MI risk</i>
-----------------	-------------------------------------

---

### Description

This dataset, `statinMIrisk_df`, is a data frame containing results from 4 clinical trials investigating the effect of statin therapy intensity on the risk of myocardial infarction or coronary death. The data compares intensive versus standard statin regimens.

### Usage

```
data(statinMIrisk_df)
```

### Format

A data frame with 4 observations and 5 variables:

**study** Study identifier (character)

**eI** Number of events in intensive treatment group (numeric)

**nI** Total patients in intensive treatment group (numeric)

**eC** Number of events in control/standard group (numeric)

**nC** Total patients in control/standard group (numeric)

### Details

The dataset name has been kept as `'statinMIrisk_df'` to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the `CardioDataSets` package and assists users in identifying its specific characteristics. The suffix `'df'` indicates that the dataset is a standard data frame. The original content has not been modified in any way.

### Source

Data taken from the RTSA package version 0.2.2

---

sulphinpyrazone_tbl_df	<i>Sulphinpyrazone for post-MI death prevention</i>
------------------------	---

---

### Description

This dataset, `sulphinpyrazone_tbl_df`, is a tibble containing information from a clinical trial studying the efficacy of sulphinpyrazone in preventing sudden death after myocardial infarction. The data includes 1,475 patients randomly assigned to either the treatment or control group.

**Usage**

```
data(sulphinpyrazone_tbl_df)
```

**Format**

A tibble with 1,475 observations and 2 variables:

**group** Treatment assignment (factor with 2 levels: "control" and "treatment")

**outcome** Patient outcome (factor with 2 levels)

**Details**

The dataset name has been kept as 'sulphinpyrazone\_tbl\_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the CardioDataSets package and assists users in identifying its specific characteristics. The suffix 'tbl\_df' indicates that the dataset is a tibble. The original content has not been modified in any way.

**Source**

Data taken from the openintro package version 2.5.0. Original study: Anturane Reinfarction Trial Research Group (1980). "Sulfinpyrazone in the prevention of sudden death after myocardial infarction." *New England Journal of Medicine*, 302(5):250-256.

---

usMortality\_df

*US Mortality Rates by Cause and Gender*

---

**Description**

This dataset, usMortality\_df, is a data frame containing mortality rates across all ages in the USA from 2011-2013, stratified by cause of death, sex, and rural/urban status. It includes national aggregate rates for 10 causes of death, including Heart disease.

**Usage**

```
data(usMortality_df)
```

**Format**

A data frame with 40 observations and 5 variables:

**Status** Residential status (factor: Rural/Urban)

**Sex** Gender (factor: Male/Female)

**Cause** Cause of death (factor with 10 levels)

**Rate** Mortality rate per 100,000 population (numeric)

**SE** Standard error of mortality rate (numeric)

## Details

The dataset name has been kept as 'usMortality\_df' to avoid confusion with other datasets in the R ecosystem. This naming convention helps distinguish this dataset as part of the CardioDataSets package and assists users in identifying its specific characteristics. The suffix 'df' indicates that the dataset is a standard data frame. The original content has not been modified in any way.

## Source

Data taken from the lattice package version 0.22-6. Original source: Rural Health Reform Policy Research Center (2015). "Exploring Rural and Urban Mortality Differences." Bethesda, MD: August 2015.

---

view\_datasets

*View Available Datasets in CardioDataSets*

---

## Description

This function lists all datasets available in the 'CardioDataSets' package. If the 'CardioDataSets' package is not loaded, it stops and shows an error message. If no datasets are available, it returns a message and an empty vector.

## Usage

```
view_datasets()
```

## Value

A character vector with the names of the available datasets. If no datasets are found, it returns an empty character vector.

## Examples

```
if (requireNamespace("CardioDataSets", quietly = TRUE)) {  
  library(CardioDataSets)  
  view_datasets()  
}
```

# Index

acs\_patients\_df, 2  
age\_heartrate\_df, 4  
ami\_occurrences\_tbl\_df, 4  
aortaDiss\_tbl\_df, 5  
  
betablockers\_matrix, 6  
  
cad\_anticoagulants\_df, 7  
cardiac\_failure\_df, 8  
cardiac\_gwas\_df, 9  
cardio\_diabetes\_tbl\_df, 12  
CardioDataSets, 10  
CardioDataSets-package  
    (CardioDataSets), 10  
cardioRiskFactors\_df, 10  
cardiovascular\_list, 11  
coronary\_death\_df, 13  
cpr\_survival\_tbl\_df, 14  
cv\_mortality\_ts, 14  
  
emotion\_heartrate\_df, 15  
  
heart\_transplant\_df, 19  
heartdisease\_tbl\_df, 17  
heartdiseaserisk\_tbl\_df, 16  
heartfailure\_df, 18  
heartTransplantTime\_tbl\_df, 19  
hfPrevention\_mtc\_network, 20  
  
mriCardioVars\_tbl\_df, 21  
muscatine\_coronary\_risk\_df, 23  
myocardialinfarction\_df, 24  
  
patient\_CAV\_df, 25  
  
radial\_ivus\_df, 25  
  
scottish\_CVD\_df, 27  
statinMIrisk\_df, 28  
sulphinpyrazone\_tbl\_df, 28  
  
usMortality\_df, 29  
  
view\_datasets, 30