

Package: tidytitanic (via r-universe)

July 7, 2026

Title Dataframes Based on Titanic Passengers and Crew

Version 0.0.1

Description The tidytitanic package provides a version of the Titanic data tailored for people analytics demonstrations and practice. Unlike other packages such as titanic, which reproduce the Kaggle competition files with minimal preprocessing, tidytitanic combines and tidies the train and test datasets into a single unified dataset, `passengers`, facilitating exploration across the full set of individuals. It also extracts meaningful personal identifiers—such as first names, last names, and titles—from the raw Name field, enabling richer demographic analysis and grouping. The `passengers` data does not cover the crew, however. Therefore the package also provides a flat and uncounted version of the `Titanic` dataset which contains crew information for further practice. This human-centered data package is designed to support exploratory data analysis, feature engineering, and pedagogical use cases.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.1

Depends R (>= 2.10)

Repository <https://evamaerey.r-universe.dev>

Date/Publication 2025-06-24 17:39:41 UTC

RemoteUrl <https://github.com/evamaerey/tidyttitanic>

RemoteRef HEAD

RemoteSha 6ea49e4e9d913e47c72f1e739bed35b135b6602b

Contents

flat_titanic	2
------------------------	---

2

passengers

passengers 2

tidy_titanic 3

Index 4

flat_titanic	<i>Tallied characteristics</i>
--------------	--------------------------------

Description

A dataset tallying frequencies of titanic passenger characteristics based on Titanic data from base R. Includes Crew

Usage

flat_titanic

Format

A data frame with 32 rows and 5 variables:

- class** class (1st, 2nd, 3rd or Crew)
- sex** sex of passenger
- age** child or adult
- survived** survival outcome of passanger
- freq** tally of all in category ...

Source

R stats

passengers	<i>Titanic test data.</i>
------------	---------------------------

Description

Titanic test data.

Usage

passengers

Format

Data frame with columns

passenger_id Passenger ID
survived survival status
pclass Passenger Class
name Name
sex Sex
age Age
sib_sp Number of Siblings/Spouses Aboard
parch Number of Parents/Children Aboard
ticket Ticket Number
fare Passenger Fare
cabin Cabin
embarked Port of Embarkation
 ...

Source

titanic package, <https://www.kaggle.com/c/titanic/data>

tidy_titanic	<i>A dataset titanic passengers and characteristics based on Titanic data from base R. Includes Crew</i>
--------------	--

Description

A dataset titanic passengers and characteristics based on Titanic data from base R. Includes Crew

Usage

```
tidy_titanic
```

Format

A data frame with 2201 rows and 5 variables:

id an id for each passenger
class class (1st, 2nd, 3rd or Crew)
sex sex of passenger
age child or adult
survived survival outcome of passanger ...

Source

R stats

Index

* datasets

flat_titanic, [2](#)

passengers, [2](#)

tidy_titanic, [3](#)

flat_titanic, [2](#)

passengers, [2](#)

tidy_titanic, [3](#)