Package 'tidytitanic'

July 22, 2025

Title Dataframes Based on Titanic Passengers and Crew

Version 0.0.1

Description

A version of the Titanic survival data tailored for people analytics demonstrations and practice. While another package, 'titanic', reproduces the Kaggle competition files with minimal preprocessing, 'tidytitanic' combines the train and test datasets into the single dataset, 'passengers', for exploration and summary across all passengers. It also extracts personal identifiers—such as first names, last names, and titles from the raw 'name' field, enabling demographic analysis. The 'passengers' data does not cover the crew, but this package also provides the more bare-bones, crew-containing datasets 'tidy_titanic' and 'flat_titanic' based on the 'Titanic' data set from 'datasets' for further exploration. 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

RoxygenNote 7.3.1

Depends R (>= 2.10)

NeedsCompilation no

Author Evangeline Reynolds [aut, cre]

Maintainer Evangeline Reynolds <evangeline.mae@gmail.com>

Repository CRAN

Date/Publication 2025-07-11 12:30:07 UTC

Contents

`flat_titanic`	*Tallied characteristics*

Description

A dataset tallying frequencies of titanic passenger characteristics based on Titanic data from datasets. 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`	*Titanic test data.*

Description

Titanic test data.

Usage

```
passengers
```

Format

Data frame with columns

passenger_id Passenger ID

title Passenger title

last_name Passenger last name

first_name Passenger first name

survived survival status

pclass Passenger Class

sex Sex

age Age

sib_sp Number of Siblings/Spouses Aboard

parch Number of Parents/Children Aboard

fare Passenger Fare

cabin Cabin

embarked Port of Embarkation

ticket Ticket Number

maiden_name Maiden name

name Full name

prefered_name Prefered name

...

Source

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

---

tidy_titanic	*A dataset titanic passengers and characteristics based on Titanic data fromdatasets. Includes Crew*

Description

A dataset titanic passengers and characteristics based on Titanic data fromdatasets. 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