

# Package ‘survivoR’

May 13, 2022

**Type** Package

**Title** Data from all Seasons of Survivor (US) TV Series in Tidy Format

**Version** 1.0

## Description

Several datasets which detail the results and events of each season of Survivor. This includes details on the cast, voting history, immunity and reward challenges, jury votes and viewers. This data is useful for practicing data wrangling, graph analytics and analysing how each season of Survivor played out.

Includes 'ggplot2' scales and colour palettes for visualisation.

**Depends** R (>= 3.5.0)

**Imports** dplyr, tidyr, ggplot2, stringr, magrittr

**Suggests** forcats, glue

**License** MIT + file LICENSE

**URL** <https://github.com/doehm/survivoR>

**BugReports** <https://github.com/doehm/survivoR/issues>

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.1.1

**NeedsCompilation** no

**Author** Daniel Oehm [aut, cre],  
Carly Levitz [ctb]

**Maintainer** Daniel Oehm <danieloehm@gmail.com>

**Repository** CRAN

**Date/Publication** 2022-05-13 06:30:02 UTC

**R topics documented:**

advantage_details . . . . .	2
advantage_movement . . . . .	3
boot_mapping . . . . .	4
castaways . . . . .	5
castaway_details . . . . .	6
challenge_description . . . . .	7
challenge_results . . . . .	8
clean_votes . . . . .	9
confessionals . . . . .	10
hidden_idols . . . . .	11
jury_votes . . . . .	12
season_palettes . . . . .	13
season_summary . . . . .	13
survivor_pal . . . . .	14
tribes_pal . . . . .	15
tribe_colours . . . . .	16
tribe_mapping . . . . .	18
viewers . . . . .	19
vote_history . . . . .	20
<b>Index</b>	<b>22</b>

---

advantage_details	<i>Advantage Details</i>
-------------------	--------------------------

---

**Description**

A dataset containing the details and characteristics of each idol and advantage. This maps to ‘advantage\_movement’

**Usage**

```
advantage_details
```

**Format**

This data frame contains the following columns:

version Country code for the version of the show

version\_season Version season key

season\_name The season name

season The season number

advantage\_id The ID / primary key of the advantage

advantage\_type Advantage type e.g. hidden immunity idol, extra vote, steal a vote, etc

clue\_details Details if a clue existed for the advantage and if so where was the clue found

location\_found The location the idol or advantage was found  
 conditions Extra details about the unique conditions of the idol or advantage

### Details

There are split idols which need to be combined to be played. In these case the first one found is given an ID. The second or subsequent parts are given the same ID with a trailing letter. For example in season 40 Denise found an idol that was split (USHI4002). Later she found the other half (USHI4002b). When played the second half is considered to have 'absorbed' into the first idol. The first idol found is always considered the primary idol.

---

advantage\_movement      *Advantage Movement*

---

### Description

A dataset containing the movement details of each advantage or hidden immunity idol. Each row is considered an event e.g. the idol was found, played, etc. If the advantage changed hands it records who received it. The logical flow is identified by the 'sequence\_id'.

### Usage

advantage\_movement

### Format

This data frame contains the following columns:

version Country code for the version of the show

version\_season Version season key

season\_name The season name

season The season number

castaway Name of the castaway involved in the event e.g. found, played, received, etc.

castaway\_id ID of the castaway (primary key). Consistent across seasons and name changes e.g. Amber Brkich / Amber Mariano. The first two letters reference the country of the version played e.g. US, AU.

advantage\_id The ID / primary key of the advantage

sequence\_id The sequence of events. For example 'sequence\_id == 1' usually means the advantage was found. Each subsequent event follows the 'sequence\_id'

day The day the event occurred

episode The episode the event occurred

event The event e.g. the advantage was found, played, received, etc

played\_for If the advantage or idol was played this records who it was played for

played\_for\_id the ID for who the advantage or idol was played for

success If the play was successful or not. Only relevant for advantages since playing a hidden immunity idol is always successful in terms of saving who it was played for.

votes\_nullified In the case of hidden immunity idols this is the count of how many votes were nullified when played

---

boot\_mapping

*Boot mapping*

---

### Description

A mapping table for easily filtering to the set of castaways that are still in the game after a specified number of boots.

### Usage

boot\_mapping

### Format

This data frame contains the following columns:

version Country code for the version of the show

version\_season Version season key

season\_name The season name

season The season number

episode Episode number

order The number of boots that there have been in the game e.g. if 'order == 2' there have been 2 boots in the game so far and there are N-2 castaways left in the game

castaway\_id ID of the castaway (primary key). Consistent across seasons and name changes e.g. Amber Brkich / Amber Mariano. The first two letters reference the country of the version played e.g. US, AU.

castaway Name of the castaway

tribe Name of the tribe the castaway was on

tribe\_status The status of the tribe e.g. original, swapped, merged, etc. See details for more

in\_the\_game Logical flag to identify if the castaway is currently in the game. If 'FALSE' the castaway is on Redemption Island or Edge of Extinction.

### Source

[https://en.wikipedia.org/wiki/Survivor\\_\(American\\_TV\\_series\)](https://en.wikipedia.org/wiki/Survivor_(American_TV_series))

---

 castaways

*Castaways*


---

### Description

A dataset containing details on the results for every castaway and season

### Usage

```
castaways
```

### Format

This data frame contains the following columns:

`version` Country code for the version of the show

`version_season` Version season key

`season` Season number

`season_name` Season name

`full_name` Full name of the castaway

`castaway_id` ID of the castaway (primary key). Consistent across seasons and name changes e.g. Amber Brkich / Amber Mariano. The first two letters reference the country of the version played e.g. US, AU (TBA).

`castaway` Name of castaway. Generally this is the name they were most commonly referred to or nickname e.g. no one called Coach, Benjamin. He was simply Coach

`age` Age of the castaway during the season they played

`city` City of residence during the season they played

`state` State of residence during the season they played

`personality_type` The Myer-Briggs personality type of the castaway. This will be removed from this table and maintained on `castaway_details` in later releases

`episode` Episode number

`day` Number of days the castaway survived. A missing value indicates they later returned to the game that season

`order` Boot order. Order in which castaway was voted out e.g. 5 is the 5th person voted of the island

`result` Final result

`jury_status` Jury status

`original_tribe` Original tribe name

### Details

If the original `castaway_id` is desired simply extract the digits from the ID e.g. `castaway_id = as.numeric(str_extract(castaway_id, '[:digit:]+'))` in a mutate step.

**Source**

[https://en.wikipedia.org/wiki/Survivor\\_\(American\\_TV\\_series\)](https://en.wikipedia.org/wiki/Survivor_(American_TV_series))

**Examples**

```
library(dplyr)
castaways %>%
  filter(season == 40)
```

---

castaway_details	<i>Castaway details</i>
------------------	-------------------------

---

**Description**

A dataset containing details on the castaways for each season

**Usage**

```
castaway_details
```

**Format**

This data frame contains the following columns:

`castaway_id` ID of the castaway (primary key). Consistent across seasons and name changes e.g. Amber Brkich / Amber Mariano. The first two letters reference the country of the version played e.g. US, AU (TBA).

`full_name` Full name of the castaway

`short_name` Short name of the castaway. Name typically used during the season. Sometimes there are multiple people with the same name e.g. Rob C and Rob M in Survivor All-Stars. This field takes the most verbose name used

`date_of_birth` Date of birth

`date_of_death` Date of death

`gender` Gender of castaway

`poc` POC indicator if known and can point to a source, else marked as white. Please log corrections on the Github page

`race` Race (if known)

`ethnicity` Ethnicity (if known)

`occupation` Occupation

`personality_type` The Myer-Briggs personality type of the castaway

**Details**

Race and ethnicity data is included if known and can point to a source, rather than making an assumption about an individual.

**Source**

[https://survivor.fandom.com/wiki/Main\\_Page](https://survivor.fandom.com/wiki/Main_Page), <https://www.personality-database.com/>

**Examples**

```
library(dplyr)
castaway_details |>
  count(gender)
```

---

challenge\_description *Challenge Description*

---

**Description**

A dataset detailing the challenges played and the elements they include over all seasons of Survivor

**Usage**

```
challenge_description
```

**Format**

This data frame contains the following columns:

challenge\_id Primary key

challenge\_name The name of the challenge. Challenges can go by different names but where possible recurring challenges are kept consistent. While there are tweaks to the challenges where the main components of the challenge consistent they share the same name

puzzle If the challenge contains a puzzle element

race If the challenge is a race between tribes, teams or individuals

precision If the challenge contains a precision element e.g. shooting an arrow, hitting a target, etc

endurance If the challenge is an endurance event e.g. last tribe, team, individual standing

strength If the challenge has a strength based

turn\_based If the challenge is turn bases i.e. conducted in rounds

balance If the challenge contains a balancing element. My refer to the player balancing on something or the player balancing an object on something e.g. The Ball Drop

food If the challenge contains a food element e.g. the food challenge, biting off chunks of meat

knowledge If the challenge contains a knowledge component e.g. Q and A about the location

memory If the challenge contains a memory element e.g. memorising a sequence of items

fire If the challenge contains an element of fire making / maintaining

water If the challenge is held, in part, in the water

**Details**

The features of each challenge have been determined largely through string searches of key words or phrases in the challenge description. It may not capture the full essence of the challenge but on the whole will provide a good basis for analysis.

Please log any suggested corrections at <https://github.com/doehm/survivoR>

For updated data please see the git version.

**Source**

<https://survivor.fandom.com/wiki/Category:Challenges>

**Examples**

```
library(dplyr)
library(tidyr)
challenge_description
```

---

challenge_results	<i>Challenge Results</i>
-------------------	--------------------------

---

**Description**

A dataset detailing the challenges played including reward and immunity challenges. immunity and rewards datasets.

**Usage**

```
challenge_results
```

**Format**

This nested data frame contains the following columns:

season\_name The season name

season The season number

episode Episode number

day The day of the tribal council

order The number of boots that there have been in the game e.g. if 'order == 2' there have been 2 boots in the game so far and there are N-2 castaways left in the game

episode\_title Episode title

challenge\_name The name of the challenge. Challenges can go by different names but where possible recurring challenges are kept consistent. While there are tweaks to the challenges where the main components of the challenge consistent they share the same name

challenge\_type The challenge type e.g. immunity, reward, etc



outcome\_type Whether the challenge is individual or tribal. Some individual reward challenges may involve multiple castaways as the winner gets to choose who they bring along  
 challenge\_id Primary key to the challenge\_description data set which contains features of the challenge  
 winning\_tribe Name of the winner tribe. NA during the merge  
 outcome\_status Identifies the winner of individual reward challenges and those chosen to participate i.e. they didn't win but were chosen by the winner to join them on the reward.  
 winner The list of winners. Either the list of people in the tribe which won, list of people that participated on the reward or the individual winner  
 winner\_id The ID of the winners of the challenge. Consistent with castaway\_id

### Details

A nested tidy data frame of immunity and reward challenge results. The winners and winning tribe of the challenge are found by expanding the winner column. For individual immunity challenges the winning tribe is simply NA.

Typically in the merge if a single person win a reward they are allowed to bring others along with them. The first castaway in the expanded list is likely to be the winner and the subsequent players those they brought along with them. Although, not always. Occasionally in the merge the castaways are split into two teams for the purpose of the reward, in which case all castaways win the reward rather than a single person.

The day field on this data set represents the day of the tribal council rather than the day of the challenge. This is to more easily associate the reward challenge with the immunity challenge and result of the tribal council. It also helps for joining tables.

### Source

[https://en.wikipedia.org/wiki/Survivor\\_\(American\\_TV\\_series\)](https://en.wikipedia.org/wiki/Survivor_(American_TV_series))

### Examples

```

library(dplyr)
library(tidyr)
challenge_results %>%
  filter(season == 40) %>%
  unnest(winners)

```

---

clean\_votes

*Cleans votes*

---

### Description

There are certain events in the game of survivor which mean someone may attend tribal council and not get the chance to vote for some reason or their vote is unique e.g. when rocks are drawn. You may want to remove the votes that were not an actual vote for a person. clean\_votes is a convenience function to remove these records. Can be piped.

**Usage**

```
clean_votes(df)
```

**Arguments**

df                    Data frame which must contain the vote data.

**Value**

Returns a tidy data frame

**Examples**

```
library(dplyr)
vh <- vote_history |>
  filter(
    season == 40,
    episode == 10
  ) |>
  count(vote)
vh

vh |>
  clean_votes()
```

---

confessionals

*Confessionals*

---

**Description**

A dataset containing the count of confessionals per castaway per episode. A confessional is when the castaway is speaking directly to the camera about their game.

**Usage**

```
confessionals
```

**Format**

This data frame contains the following columns:

version    Country code for the version of the show

version\_season    Version season key

season\_name    The season name

season    The season number

episode    Episode number

castaway    Name of the castaway

castaway\_id ID of the castaway (primary key). Consistent across seasons and name changes e.g. Amber Brkich / Amber Mariano. The first two letters reference the country of the version played e.g. US, AU.

confessional\_count The count of confessionals for the castaway during the episode

## Details

Confessional data has been counted by contributors of the survivoR R package and consolidated with external sources. The aim is to establish consistency in confessional counts in the absence of official sources. Given the subjective nature of the counts and the potential for clerical error no single source is more valid than another. Therefore, it is reasonable to average across all sources.

In the case of double or extended episodes, if the episode only has one title it is considered a single episode. This means the average number of confessionals per person is likely to be higher for this episode given it's length. If there are two episode titles the confessionals are counted for the appropriate episode. This is to ensure consistency across all other datasets.

In the case of recap episodes, this episode is left blank.

If you also count confessionals, please get in touch and I'll add them into the package.

---

hidden\_idols *(Deprecated) Hidden Immunity Idols*

---

## Description

This data set has been replaced with the advantage data sets which contain more details and easier to use structure

## Usage

hidden\_idols

## Format

This data frame contains the following columns:

season\_name The season name

season The season number

castaway\_id ID of the castaway (primary key). Consistent across seasons and name changes e.g. Amber Brkich / Amber Mariano. The first two letters reference the country of the version played e.g. US, AU.

castaway Name of the castaway

idol\_number Indicates whether it is the first, second, etc idol found in the season

idols\_held The number of idols held by the castaway

votes\_nullified The number of votes nullified by the idol

day\_found The day the idol was found

day\_played The day of the tribal council

legacy\_advantage If the idol was a legacy advantage or not

**Source**

[https://survivor.fandom.com/wiki/Hidden\\_Immunity\\_Idol](https://survivor.fandom.com/wiki/Hidden_Immunity_Idol)

---

jury_votes	<i>Jury votes</i>
------------	-------------------

---

**Description**

A dataset containing details on the final jury votes to determine the winner for each season

**Usage**

jury\_votes

**Format**

This data frame contains the following columns:

version Country code for the version of the show

version\_season Version season key

season\_name The season name

season The season number

castaway Name of the castaway

finalist The finalists for which a vote can be placed

vote Vote. 0-1 variable for easy summation

castaway\_id ID of the castaway (primary key). Consistent across seasons and name changes e.g. Amber Brkich / Amber Mariano. The first two letters reference the country of the version played e.g. US, AU.

finalist\_id The ID of the finalist for which a vote can be placed. Consistent with castaway ID

**Source**

[https://en.wikipedia.org/wiki/Survivor\\_\(American\\_TV\\_series\)](https://en.wikipedia.org/wiki/Survivor_(American_TV_series))

**Examples**

```
library(dplyr)
jury_votes %>%
  filter(season == 40) %>%
  group_by(finalist) %>%
  summarise(votes = sum(vote))
```

---

season_palettes	<i>Season palettes</i>
-----------------	------------------------

---

**Description**

A dataset containing palettes generated from the season logos

**Usage**

season\_palettes

**Format**

This nested data frame contains the following columns:

season\_name The season name

season The season number

palette The season palette

**Source**

[https://en.wikipedia.org/wiki/Survivor\\_\(American\\_TV\\_series\)](https://en.wikipedia.org/wiki/Survivor_(American_TV_series))

---

season_summary	<i>Season summary</i>
----------------	-----------------------

---

**Description**

A dataset containing a summary of all 40 seasons of Survivor

**Usage**

season\_summary

**Format**

This data frame contains the following columns:

version Country code for the version of the show

version\_season Version season key

season\_name Season name

season Season number

location Location of the season

country Country the season was held

tribe\_setup Initial setup of the tribe e.g. heroes vs Healers vs Hustlers  
 full\_name Full name of the winner  
 winner\_id ID for the winner of the season (primary key)  
 winner Winner of the season  
 runner\_ups Runner ups for the season. Either one or two runner ups as a string  
 final\_vote Final vote allocation. See the jury\_votes dataset for better aggregation of this data  
 timeslot Timeslot of the show in the US  
 premiered Date the first episode aired  
 ended Date the season ended  
 filming\_started Date the filming of the season started  
 filming\_ended Date the filming ended (39 or 42 days after the start)  
 viewers\_premier Number of viewers (millions) who tuned in for the premier  
 viewers\_finale Number of viewers (millions) who tuned in for the finale  
 viewers\_reunion Number of viewers (millions) who tuned in for the reunion  
 viewers\_mean Average number of viewers (millions) who tuned in over the season  
 rank Season rank

### Source

[https://en.wikipedia.org/wiki/Survivor\\_\(American\\_TV\\_series\)](https://en.wikipedia.org/wiki/Survivor_(American_TV_series))

---

survivor_pal	<i>Survivor season colour palette</i>
--------------	---------------------------------------

---

### Description

ggplot2 scales for each season of Survivor.

### Usage

```

survivor_pal(season = NULL, scale_type = "d", reverse = FALSE, ...)
scale_fill_survivor(season = NULL, scale_type = "d", reverse = FALSE, ...)
scale_colour_survivor(season = NULL, scale_type = "d", reverse = FALSE, ...)

```

### Arguments

season	Season number
scale_type	Discrete or continuous. Input d or c.
reverse	Logical. Reverse the palette?
...	Other arguments passed on to methods.

**Details**

Palettes are created from the logo for the season.

**Value**

Scale functions for ggplot2

**Examples**

```
library(ggplot2)
library(dplyr)
mpg %>%
  ggplot(aes(x = displ, fill = manufacturer)) +
  geom_histogram(colour = "black") +
  scale_fill_survivor(40)
```

---

tribes\_pal

*Tribes colour palette*


---

**Description**

To create scale functions for ggplot. Given a season of Survivor, a palette is created from the tribe colours for that season including the merged tribe.

**Usage**

```
tribes_pal(season = NULL, scale_type = "d", reverse = FALSE, tribe = NULL, ...)
```

```
scale_fill_tribes(season = NULL, scale_type = "d", reverse = FALSE, ...)
```

```
scale_colour_tribes(season = NULL, scale_type = "d", reverse = FALSE, ...)
```

**Arguments**

season	Season number
scale_type	Discrete or continuous. Input d or c.
reverse	Logical. Reverse the palette?
tribe	Tribe names. Default NULL
...	Other arguments passed on to methods.

**Details**

If it is intended the colours will correspond to the tribes e.g. a stacked bar chart of votes given to each finalist and the colour corresponds to their original tribe (as in the example below), the tribe vector needs to be passed to the scale function (for now). If no tribe vector is given it will simply treat the tribe colours as a colour palette.

**Value**

Scale functions for ggplot2

**Examples**

```
library(ggplot2)
library(stringr)
library(dplyr)
library(glue)
ssn <- 35
labels <- castaways %>%
  filter(
    season == ssn,
    str_detect(result, "Sole|unner")
  ) %>%
  select(castaway, original_tribe) %>%
  mutate(label = glue("{castaway} ({original_tribe}")) %>%
  select(label, castaway)
jury_votes %>%
  filter(season == ssn) %>%
  left_join(
    castaways %>%
      filter(season == ssn) %>%
      select(castaway, original_tribe),
    by = "castaway"
  ) %>%
  group_by(finalist, original_tribe) %>%
  summarise(votes = sum(vote)) %>%
  left_join(labels, by = c("finalist" = "castaway")) %>% {
  ggplot(., aes(x = label, y = votes, fill = original_tribe)) +
  geom_bar(stat = "identity", width = 0.5) +
  scale_fill_tribes(ssn, tribe = .$original_tribe) +
  theme_minimal() +
  labs(
    x = "Finalist (original tribe)",
    y = "Votes",
    fill = "Original\\ntribe",
    title = "Votes received by each finalist"
  )
}
```

---

tribe\_colours

*Tribe colours*

---

**Description**

A dataset containing the tribe colours for each season

**Usage**

tribe\_colours



**Format**

This data frame contains the following columns:

version Country code for the version of the show

version\_season Version season key

season\_name The season name

season The season number

tribe Tribe name

tribe\_colour Colour of the tribe

tribe\_status Tribe status e.g. original, swapped or merged. In the instance where a tribe is formed at the swap by splitting 2 tribes into 3, the 3rd tribe will be labelled 'swapped'

**Source**

<https://survivor.fandom.com/wiki/Tribe>

**Examples**

```
library(ggplot2)
library(dplyr)
library(forcats)
df <- tribe_colours %>%
  group_by(season_name) %>%
  mutate(
    xmin = 1,
    xmax = 2,
    ymin = 1:n(),
    ymax = ymin + 1
  ) %>%
  ungroup() %>%
  mutate(
    season_name = fct_reorder(season_name, season),
    font_colour = ifelse(tribe_colour == "#000000", "white", "black")
  )
ggplot() +
  geom_rect(data = df,
    mapping = aes(xmin = xmin, xmax = xmax, ymin = ymin, ymax = ymax),
    fill = df$tribe_colour) +
  geom_text(data = df,
    mapping = aes(x = xmin+0.5, y = ymin+0.5, label = tribe),
    colour = df$font_colour) +
  theme_void() +
  facet_wrap(~season_name, scales = "free_y")
```

---

tribe_mapping	<i>Tribe mapping</i>
---------------	----------------------

---

### Description

A mapping for castaways to tribes for each day (day being the day of the tribal council) This is useful for observing who is on what tribe throughout the game.

### Usage

```
tribe_mapping
```

### Format

This data frame contains the following columns:

version Country code for the version of the show

version\_season Version season key

season\_name The season name

season The season number

episode Episode number

day The day of the tribal council

castaway\_id ID of the castaway (primary key). Consistent across seasons and name changes e.g. Amber Brkich / Amber Mariano. The first two letters reference the country of the version played e.g. US, AU.

castaway Name of the castaway

tribe Name of the tribe the castaway was on

tribe\_status The status of the tribe e.g. original, swapped, merged, etc. See details for more

### Details

Each season by episode and day holds a complete list of castaways still in the game and which tribe they are on. Moving through each day you can observe the changes in the tribe. For example the first day has all castaways mapped to their original tribe. The next day has the same minus the castaway just voted out. This is useful for observing the changes in tribe make either due to castaways being voted off the island, tribe swaps, who is on Redemption Island and Edge of Extinction.

### Source

[https://en.wikipedia.org/wiki/Survivor\\_\(American\\_TV\\_series\)](https://en.wikipedia.org/wiki/Survivor_(American_TV_series))

---

viewers	<i>Viewers</i>
---------	----------------

---

**Description**

A dataset containing the viewer history for each season and episode

**Usage**

viewers

**Format**

This data frame contains the following columns:

version Country code for the version of the show

version\_season Version season key

season\_name The season name

season Season number

episode\_number\_overall The cumulative episode number

episode Episode number for the season

episode\_title Episode title

episode\_date Date the episode aired

viewers Number of viewers (millions) who tuned in

rating\_18\_49 TV rating for the 18-49 aged group

share\_18\_49 TV share for the 18-49 aged group

imdb\_rating IMDB rating for the episode on a scale of 0-10

**Source**

[https://en.wikipedia.org/wiki/Survivor\\_\(American\\_TV\\_series\)](https://en.wikipedia.org/wiki/Survivor_(American_TV_series))

---

vote_history	<i>Vote history</i>
--------------	---------------------

---

### Description

A dataset containing details on the vote history for each season

### Usage

vote\_history

### Format

This data frame contains the following columns:

version Country code for the version of the show

version\_season Version season key

season\_name The season name

season The season number

episode Episode number

day Day the tribal council took place

tribe\_status The status of the tribe e.g. original, swapped, merged, etc. See details for more

tribe Tribe name

castaway Name of the castaway

immunity Type of immunity held by the castaway at the time of the vote e.g. individual, hidden (see details for hidden immunity data)

vote The castaway for which the vote was cast

vote\_event Extra details on the vote e.g. Won or lost the fire challenge, played an extra vote, etc

split\_vote If there was a decision to split the vote this records who the vote was split with. Helps to identify successful boots

nullified Was the vote nullified by a hidden immunity idol? Logical

tie If the set of votes resulted in a tie. Logical

voted\_out The castaway who was voted out

order Boot order. Order in which castaway was voted out e.g. 5 is the 5th person voted of the island

vote\_order In the case of ties this indicates the order the votes took place

castaway\_id ID of the castaway (primary key). Consistent across seasons and name changes e.g. Amber Brkich / Amber Mariano. The first two letters reference the country of the version played e.g. US, AU.

vote\_id ID of the castaway voted for

voted\_out\_id ID of the castaway voted\_out

## Details

This data frame contains a complete history of votes cast across all seasons of Survivor. While there are consistent events across the seasons there are some unique events such as the 'mutiny' in Survivor: Cook Islands (season 13) or the 'Outcasts' in Survivor: Pearl Islands (season 7). For maintaining a standard, whenever there has been a change in tribe for the castaways it has been recorded as swapped. swapped is used as the term since 'the tribe swap' is a typical recurring milestone in each season of Survivor. Subsequent changes are recorded with a trailing digit e.g. swapped2. This includes absorbed tribes e.g. Stephanie was 'absorbed' in Survivor: Palau (season 10) and when 3 tribes are reduced to 2. These cases are still considered 'swapped' to indicate a change in tribe status.

Some events result in a castaway attending tribal but not voting. These are recorded as

Win The castaway won the fire challenge

Lose The castaway lost the fire challenge

None The castaway did not cast a vote. This may be due to a vote steal or some other means

Immune The castaway did not vote but were immune from the vote

Where a castaway has immunity == 'hidden' this means that player is protected by a hidden immunity idol. It may not necessarily mean they played the idol, the idol may have been played for them. While the nullified votes data is complete the immunity data does not include those who had immunity but did not receive a vote. This is a TODO.

In the case where the 'steal a vote' advantage was played, there is a second row for the castaway that stole the vote. The castaway who had their vote stolen are is recorded as None.

Many castaways have been medically evacuated, quit or left the game for some other reason. In these cases where no votes were cast there is a skip in the order variable. Since no votes were cast there is nothing to record on this data frame. The correct order in which castaways departed the island is recorded on castaways.

In the case of a tie, voted\_out is recorded as tie to indicate no one was voted off the island in that instance. The re-vote is recorded with vote\_order = 2 to indicate this is the second round of voting. In the case of a second tie voted\_out is recorded as tie2. The third step is either a draw of rocks, fire challenge or countback (in the early days of survivor). In these cases vote is recorded as the colour of the rock drawn, result of the fire challenge or 'countback'.

## Source

[https://en.wikipedia.org/wiki/Survivor\\_\(American\\_TV\\_series\)](https://en.wikipedia.org/wiki/Survivor_(American_TV_series))

## Examples

```
# The number of times Tony voted for each castaway in Survivor: Winners at War
library(dplyr)
vote_history %>%
  filter(
    season == 40,
    castaway == "Tony"
  ) %>%
  count(vote)
```

# Index

## \* datasets

- advantage\_details, 2
  - advantage\_movement, 3
  - boot\_mapping, 4
  - castaway\_details, 6
  - castaways, 5
  - challenge\_description, 7
  - challenge\_results, 8
  - confessionals, 10
  - hidden\_idols, 11
  - jury\_votes, 12
  - season\_palettes, 13
  - season\_summary, 13
  - tribe\_colours, 16
  - tribe\_mapping, 18
  - viewers, 19
  - vote\_history, 20
- advantage\_details, 2
- advantage\_movement, 3
- boot\_mapping, 4
- castaway\_details, 6
- castaways, 5
- challenge\_description, 7
- challenge\_results, 8
- clean\_votes, 9
- confessionals, 10
- hidden\_idols, 11
- jury\_votes, 12
- scale\_colour\_survivor (survivor\_pal), 14
- scale\_colour\_tribes (tribes\_pal), 15
- scale\_fill\_survivor (survivor\_pal), 14
- scale\_fill\_tribes (tribes\_pal), 15
- season\_palettes, 13
- season\_summary, 13
- survivor\_pal, 14
- tribe\_colours, 16
- tribe\_mapping, 18
- tribes\_pal, 15
- viewers, 19
- vote\_history, 20