

Package ‘gendercoder’

May 18, 2026

Title Recodes Sex/Gender Descriptions into a Standard Set

Version 0.1.1

Description Provides dictionary-based tools for recoding free-text gender responses into consistent categories while preserving gender diversity where possible. The package standardises spelling, capitalization, whitespace, and common variants through curated named character-vector dictionaries, supports either detailed or collapsed output categories, and can retain original unmatched responses for manual review. It also includes helpers for creating custom dictionaries from approximate string matches and a local interactive application for recoding uploaded data files.

Depends R (>= 3.0.0)

Maintainer Yaoxiang Li <liyaoxiang@outlook.com>

License GPL-2

Encoding UTF-8

LazyData true

RoxygenNote 7.3.3

Suggests knitr, rmarkdown, bs4Dash, haven, shiny

VignetteBuilder knitr

URL <https://github.com/ropensci/gendercoder>

BugReports <https://github.com/ropensci/gendercoder/issues>

NeedsCompilation no

Author Yaoxiang Li [aut, cre] (ORCID: <<https://orcid.org/0000-0001-9200-1016>>),
Jennifer Beaudry [aut] (ORCID: <<https://orcid.org/0000-0003-1596-6708>>),
Emily Kothe [aut] (ORCID: <<https://orcid.org/0000-0003-1210-0554>>),
Felix Singleton Thorn [aut] (ORCID:
<<https://orcid.org/0000-0002-0237-6146>>),
Rhydwyn McGuire [aut],
Nicholas Tierney [aut] (ORCID: <<https://orcid.org/0000-0003-1460-8722>>),
Mathew Ling [aut] (ORCID: <<https://orcid.org/0000-0002-0940-2538>>),
Julia Silge [rev] (Julia reviewed the package (v. 0.0.0.9000) for
rOpenSci, see

<<https://github.com/ropensci/software-review/issues/435>>),
 Elin Waring [rev] (Elin reviewed the package (v. 0.0.0.9000) for
 rOpenSci, see
 <<https://github.com/ropensci/software-review/issues/435>>)

Repository CRAN

Date/Publication 2026-05-18 18:10:16 UTC

Contents

fewlevels_en	2
gendercoder	2
gendercoder_app	3
gender_create_dictionary	4
manylevels_en	4
recode_gender	5
sample	6
Index	7

fewlevels_en	<i>fewlevels_en</i>
--------------	---------------------

Description

An English dictionary for the recode_gender function that has fewer levels

gendercoder	<i>gendercoder: A Package for Recoding Freetext Gender Data</i>
-------------	---

Description

Provides dictionaries and recode_gender() to allow for easy automatic coding of common variations in free-text responses to the question **"What is your gender?"**

Author(s)

Maintainer: Yaoxiang Li <liyaoxiang@outlook.com> ([ORCID](#))

Authors:

- Jennifer Beaudry <jbeaudry@swin.edu.au> ([ORCID](#))
- Emily Kothe <emily.kothe@deakin.edu.au> ([ORCID](#))
- Felix Singleton Thorn <fsingletonthorn@gmail.com> ([ORCID](#))
- Rhydwyn McGuire <rhydwyn@rhydwyn.net>
- Nicholas Tierney <nicholas.tierney@gmail.com> ([ORCID](#))

- Mathew Ling <mathewtyling@gmail.com> ([ORCID](#))

Other contributors:

- Julia Silge (Julia reviewed the package (v. 0.0.0.9000) for rOpenSci, see <<https://github.com/ropensci/software-review/issues/435>>) [reviewer]
- Elin Waring (Elin reviewed the package (v. 0.0.0.9000) for rOpenSci, see <<https://github.com/ropensci/software-review/issues/435>>) [reviewer]

See Also

Useful links:

- <https://github.com/ropensci/gendercoder>
- Report bugs at <https://github.com/ropensci/gendercoder/issues>

gendercoder_app

Launch the gendercoder Shiny app

Description

Code data interactively in a Shiny app that runs locally in RStudio or a web browser using a **bs4Dash** interface. The app supports CSV, Stata, SPSS, RDS, and R data files. Stata and SPSS files require the optional **haven** package.

Usage

```
gendercoder_app(...)
```

Arguments

```
...           arguments to pass to shiny::runApp()
```

Value

Called for its side effect of launching a Shiny app.

Examples

```
if (interactive()) {  
  gendercoder_app()  
}
```

gender_create_dictionary

Create a custom dictionary from fuzzy matches

Description

gender_create_dictionary suggests dictionary entries for gender responses that are not already matched exactly. The returned named character vector is intended to be reviewed before it is combined with a built-in dictionary and passed to recode_gender().

Usage

```
gender_create_dictionary(
  gender,
  dictionary = gendercoder::manylevels_en,
  max_distance = 1
)
```

Arguments

gender	a character vector of gender responses for recoding
dictionary	a character vector whose names are known gender responses and whose values are replacement values
max_distance	maximum edit distance allowed for a suggested match

Value

a named character vector of suggested replacement values

Examples

```
suggested <- gender_create_dictionary(
  c("maile", "unknown"),
  dictionary = manylevels_en,
  max_distance = 1
)
suggested
```

manylevels_en

manylevels_en

Description

An English dictionary for the recode_gender function that has many levels

recode_gender	<i>recode_gender</i>
---------------	----------------------

Description

recode_gender matches uncleaned gender responses to cleaned list using an built-in or custom dictionary.

Usage

```
recode_gender(  
  gender,  
  dictionary = gendercoder::manylevels_en,  
  retain_unmatched = FALSE  
)
```

Arguments

gender	a character vector of gender responses for recoding
dictionary	a list that the contains gender responses and their replacement values. A built-in dictionary manylevels_en is used by default if an alternative dictionary is not supplied.
retain_unmatched	logical indicating if gender responses that are not found in dictionary should be filled with the uncleaned values during recoding

Value

a character vector of recoded genders

Examples

```
df <- data.frame(  
  stringsAsFactors = FALSE,  
  gender = c("male", "MALE", "mle", "I am male", "femal", "female", "enby"),  
  age = c(34L, 37L, 77L, 52L, 68L, 67L, 83L)  
)  
  
df$recoded_gender <- recode_gender(df$gender,  
  dictionary = manylevels_en,  
  retain_unmatched = TRUE  
)  
df
```

sample

sample

Description

A sample data.frame of free-text gender in English for testing and demonstration

Index

`fewlevels_en`, [2](#)

`gender_create_dictionary`, [4](#)

`gendercoder`, [2](#)

`gendercoder-package (gendercoder)`, [2](#)

`gendercoder_app`, [3](#)

`manylevels_en`, [4](#)

`recode_gender`, [5](#)

`sample`, [6](#)