

Package ‘summarytbl’

October 6, 2025

Type Package

Title Generate Summary Tables for Continuous, Ordinal, and Categorical Data

Version 0.1.0

Maintainer Ama Nyame-Mensah <ama@anyamemensah.com>

URL <https://anyamemensah.github.io/summarytbl/>

Description Provides functions for tabulating and summarizing continuous, ordinal, and categorical variables in data frames. The package was designed to streamline exploratory data analysis and simplify the creation of summary tables for reports and other purposes.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Imports dplyr, purrr, rlang, stats, tibble, tidyr

RoxygenNote 7.3.3

Suggests knitr, rmarkdown, testthat (>= 3.0.0)

VignetteBuilder knitr

Depends R (>= 4.1.0)

NeedsCompilation no

Author Ama Nyame-Mensah [aut, cre]

Repository CRAN

Date/Publication 2025-10-06 08:00:02 UTC

Contents

cat_group_tbl	2
cat_tbl	3
check_named_vctr	4
depressive	5
mean_group_tbl	6

mean_tbl	8
nlsy	10
select_group_tbl	11
select_tbl	14
social_psy_data	16
stem_social_psy	17
tas	19

Index	21
--------------	-----------

cat_group_tbl	<i>Summarize a categorical variable by a grouping variable</i>
---------------	--

Description

cat_group_tbl() presents frequency counts and percentages (count, percent) for nominal or categorical variables by some grouping variable. Relative frequencies and percentages of each level of the primary categorical variable (row_var) within each level of the grouping variable (col_var) can be returned. Missing data can be excluded for either variable from the calculations. By default, the table is returned in the long format.

Usage

```
cat_group_tbl(
  data,
  row_var,
  col_var,
  na.rm.row_var = FALSE,
  na.rm.col_var = FALSE,
  only = NULL,
  ignore = NULL,
  pivot = "longer"
)
```

Arguments

data	A data frame.
row_var	A character string of the name of a column in data containing categorical data. This is the primary categorical variable. When pivoted to the wider format, the categories of this variable will appear in the rows of the table.
col_var	A character string of the name of a column in data containing categorical data. This is the primary categorical variable. When pivoted to the wider format, the categories of this variable will appear in the rows of the table.
na.rm.row_var	A logical value indicating whether missing values for row_var should be removed before calculations. Default is FALSE.
na.rm.col_var	A logical value indicating whether missing values for col_var should be removed before calculations. Default is FALSE.

only	A character string or vector of strings indicating the types of summary data to return. The default is NULL, which includes both counts and percentages. To return only one type, specify count or percent. Percentages are calculated column-wise, grouped by col_var.
ignore	A named character vector or list containing values to ignore from row_var and col_var.
pivot	A character string specifying the format of the returned summary table. The default is longer, which returns the data in long format. To return the data in wide format, use wider.

Value

A tibble displaying relative frequency counts and/or percentages of row_var, grouped by col_var. When the output is in wider format, columns prefixed with count_ and percent_ contain the frequency and proportion, respectively, for each distinct response value of row_var within each level of col_var.

Author(s)

Ama Nyame-Mensah

Examples

```
cat_group_tbl(data = nlsy,
              row_var = "gender",
              col_var = "bthwht",
              pivot = "wider",
              only = "count")

cat_group_tbl(data = nlsy,
              row_var = "birthord",
              col_var = "breastfed",
              pivot = "longer")
```

cat_tbl

Summarize a categorical variable

Description

cat_group_tbl() presents frequency counts and percentages (count, percent) for nominal or categorical variables. Missing data can be excluded from the calculations.

Usage

```
cat_tbl(data, var, na.rm = FALSE, only = NULL, ignore = NULL)
```

Arguments

data	A data frame.
var	A character string of the name of a variable in data containing categorical data.
na.rm	A logical value indicating whether missing values should be removed before calculations. Default is FALSE.
only	A character string, or vector of character strings, of the types of summary data to return. Default is NULL, which returns both counts and percentages. To return only counts or percentages, use count or percent, respectively.
ignore	An optional vector that contains values to exclude from the data. Default is NULL, which includes all present values.

Value

A tibble displaying the relative frequency counts and/or percentages of row_var.

Author(s)

Ama Nyame-Mensah

Examples

```
cat_tbl(data = nlsy, var = "gender")

cat_tbl(data = nlsy, var = "race", only = "count")

cat_tbl(data = nlsy,
        var = "race",
        ignore = "Hispanic",
        only = "percent",
        na.rm = TRUE)
```

check_named_vctr	<i>Check a named vector</i>
------------------	-----------------------------

Description

This function assesses whether named lists and vectors have invalid values (like NULL or NA), invalid names (such as missing or empty names), confirms that the count of valid names matches the count of provided values, and verifies that the valid names obtained from the named object align with the supplied names. If any checks fail, the default value is returned.

Usage

```
check_named_vctr(x, names, default)
```

Arguments

x A named vector.
names A character vector specifying the names to be matched.
default Default value to return

Value

Either the original object, `x`, or the `default` value.

Author(s)

Ama Nyame-Mensah

Examples

```
# returns NULL
check_named_vctr(x = c(one = 1, two = 2, 3),
                 names = c("one", "two", "three"),
                 default = NULL)

# returns x
check_named_vctr(x = list(one = 1, two = 2, three = 3),
                 names = list("one", "two", "three"),
                 default = NULL)
```

 depressive

Depressive Symptoms Data

Description

These data are a subset from the National Longitudinal Survey of Youth (NLSY) 1979 Children and Young Adults. The dataset includes information about depressive symptoms in children and young adults. The dataset has 11,551 observations and 12 variables.

For more information about the National Longitudinal Survey of Youth, visit <https://www.nlsinfo.org/>.

Usage

```
depressive
```

Format

A data.frame with 11,551 rows and 12 columns:

cid Child identification number)

race race of child (1 = Hispanic, 2 = Black, 3 = Non-Black,Non-Hispanic)

gender gender of child (1 = male, 2 = female)

- job** year of child's bith
- dep_1** how often child feels sad and blue (1 = often, 2 = sometimes, 3 = hardly ever)
- dep_2** how often child feels nervous, tense, or on edge (1 = often, 2 = sometimes, 3 = hardly ever)
- dep_3** how often child feels happy (1 = often, 2 = sometimes, 3 = hardly ever)
- dep_4** how often child feels bored (1 = often, 2 = sometimes, 3 = hardly ever)
- dep_5** how often child feels lonely (1 = often, 2 = sometimes, 3 = hardly ever)
- dep_6** how often child feels tired or worn out (1 = often, 2 = sometimes, 3 = hardly ever)
- dep_7** how often child feels excited about something (1 = often, 2 = sometimes, 3 = hardly ever)
- dep_8** how often child feels too busy to get everything (1 = often, 2 = sometimes, 3 = hardly ever)

mean_group_tbl *Summarize continuous variables by group*

Description

mean_group_tbl() presents descriptive statistics (mean, sd, minimum, maximum, number of non-missing observations) for interval (e.g., Test scores) and ratio level (e.g., Age) variables with the same variable stem by some grouping variable. A variable stem is a common prefix found in related variable names, often corresponding to similar survey items, that represents a shared concept before unique identifiers (like time points) are added. For example, in the stem_social_psych dataset, the two variables 'belong_belongStem_w1' and 'belong_belongStem_w2' share the variable stem 'belong_belongStem' (e.g., "I feel like an outsider in STEM"), with suffixes (_w1, _w2) indicating different measurement waves. By default, missing data are excluded from the calculations in a listwise fashion.

Usage

```
mean_group_tbl(
  data,
  var_stem,
  group,
  escape_stem = FALSE,
  ignore_stem_case = FALSE,
  group_type = "variable",
  group_name = NULL,
  escape_group = FALSE,
  ignore_group_case = FALSE,
  remove_group_non_alnum = TRUE,
  na_removal = "listwise",
  only = NULL,
  var_labels = NULL,
  ignore = NULL
)
```

Arguments

<code>data</code>	A data frame.
<code>var_stem</code>	A character string of a variable stem or the full name of a variable in data.
<code>group</code>	A character string of a variable in data or a pattern to use to search for variables in data.
<code>escape_stem</code>	A logical value indicating whether to escape <code>var_stem</code> . Default is FALSE.
<code>ignore_stem_case</code>	A logical value indicating whether the search for columns matching the supplied <code>var_stem</code> is case-insensitive. Default is FALSE.
<code>group_type</code>	A character string that defines the type of grouping variable. Should be one of <code>pattern</code> or <code>variable</code> . Default is <code>variable</code> , in which case the variable matching the group string will be searched for within data.
<code>group_name</code>	A character string piped to the final table to replace the name of group.
<code>escape_group</code>	A logical value indicating whether to escape string supplied to <code>group</code> .
<code>ignore_group_case</code>	A logical value indicating whether <code>group</code> is case-insensitive. Default is FALSE.
<code>remove_group_non_alnum</code>	A logical value indicating whether to remove all non- alphanumeric characters (anything that is not a letter or number) from <code>group</code> . Default is TRUE.
<code>na_removal</code>	A character string specifying how to remove missing values. Should be one of <code>pairwise</code> or <code>listwise</code> . Default is <code>listwise</code> .
<code>only</code>	A character string or vector of character strings of the types of summary data to return. Default is NULL, which returns mean (<code>mean</code>), standard deviation (<code>sd</code>), minimum value (<code>min</code>), maximum value (<code>max</code>), and non-missing responses (<code>nobs</code>).
<code>var_labels</code>	An optional named character vector or list where each element maps labels to variable names. If any element is unnamed or if any labels do not match variables in returned from data, all labels will be ignored and the table will be printed without them.
<code>ignore</code>	An optional named vector or list specifying values to exclude from the dataset and analysis. By default, NULL includes all available values. To omit values from variables returned by <code>var_stem</code> , use the provided stem as the name. To exclude values from both <code>var_stem</code> variables and a grouping variable in data, supply a list.

Value

A tibble presenting summary statistics (e.g., mean, standard deviation, minimum value, maximum, number of non-missing observations) for a set of variables sharing the same variable stem. The results are grouped by either a grouping variable in the data or by a pattern matched with variable names.

Author(s)

Ama Nyame-Mensah

Examples

```
mean_group_tbl(data = stem_social_psych,
               var_stem = "belong_welcomedStem",
               group = "_w\\d",
               group_type = "pattern",
               na_removal = "pairwise",
               var_labels = c(belong_welcomedStem_w1 = "I feel welcomed in STEM workplaces",
                             belong_welcomedStem_w2 = "I feel welcomed in STEM workplaces"),
               group_name = "wave")

mean_group_tbl(data = social_psy_data,
               var_stem = "belong",
               group = "gender",
               group_type = "variable",
               na_removal = "pairwise",
               var_labels = c(belong_1 = "I feel like I belong at this institution",
                             belong_2 = "I feel like part of the community",
                             belong_3 = "I feel valued by this institution"),
               group_name = "gender_identity")

grouped_data <-
  data.frame(
    symptoms.t1 = sample(c(1:5, -999), replace = TRUE, size = 50),
    symptoms.t2 = sample(c(NA, 1:5, -999), replace = TRUE, size = 50)
  )

mean_group_tbl(data = grouped_data,
               var_stem = "symptoms",
               group = ".t\\d",
               group_type = "pattern",
               escape_group = TRUE,
               na_removal = "listwise",
               ignore = c(symptoms = -999))
```

mean_tbl

Summarize continuous variables

Description

`mean_tbl()` presents descriptive statistics (mean, sd, minimum, maximum, number of non-missing observations) for interval (e.g., Test scores) and ratio level (e.g., Age) variables with the same variable stem. A variable stem is a common prefix found in related variable names, often corresponding to similar survey items, that represents a shared concept before unique identifiers (like time points) are added. For example, in the `stem_social_psych` dataset, the two variables `'belong_belongStem_w1'` and `'belong_belongStem_w2'` share the variable stem `'belong_belongStem'` (e.g., "I feel like an outsider in STEM"), with suffixes (`_w1`, `_w2`) indicating different measurement waves. By default, missing data are excluded from the calculations in a listwise fashion.

Usage

```
mean_tbl(  
  data,  
  var_stem,  
  escape_stem = FALSE,  
  ignore_stem_case = FALSE,  
  na_removal = "listwise",  
  only = NULL,  
  var_labels = NULL,  
  ignore = NULL  
)
```

Arguments

<code>data</code>	A data frame.
<code>var_stem</code>	A character string of a variable stem or the full name of a variable in data.
<code>escape_stem</code>	A logical value indicating whether to escape <code>var_stem</code> . Default is FALSE.
<code>ignore_stem_case</code>	A logical value indicating whether the search for columns matching the supplied <code>var_stem</code> is case-insensitive. Default is FALSE.
<code>na_removal</code>	A character string specifying how to remove missing values. Should be one of <code>pairwise</code> or <code>listwise</code> . Default is <code>listwise</code> .
<code>only</code>	A character string or vector of character strings of the kinds of summary statistics to return. Default is NULL, which returns mean (<code>mean</code>), standard deviation (<code>sd</code>), minimum value (<code>min</code>), maximum value (<code>max</code>), and non-missing responses (<code>nobs</code>).
<code>var_labels</code>	An optional named character vector or list where each element maps labels to variable names. If any element is unnamed or if any labels do not match variables in returned from data, all labels will be ignored and the table will be printed without them.
<code>ignore</code>	An optional vector that contains values to exclude from the data. Default is NULL, which includes all present values.

Value

A tibble presenting summary statistics for series of continuous variables with the same variable stem.

Author(s)

Ama Nyame-Mensah

Examples

```
mean_tbl(data = social_psy_data,  
         var_stem = "belong")
```

```
mean_tbl(data = social_psy_data,
         var_stem = "belong",
         na_removal = "pairwise",
         var_labels = c(belong_1 = "I feel like I belong at this institution",
                       belong_2 = "I feel like part of the community",
                       belong_3 = "I feel valued by this institution"))
```

nlsy

National Longitudinal Survey of Youth (NLSY) Data

Description

These data are a subset from the National Longitudinal Survey of Youth (NLSY) 1979 Children and Young Adults. The data contains 2,976 observations and 10 variables.

For more information about the National Longitudinal Survey of Youth, visit <https://www.nlsinfo.org/>.

Usage

```
nlsy
```

Format

A tibble with 2,976 rows and 11 columns:

CID Child identification number)

race race of child (Hispanic, Black, Non-Black, Non-Hispanic)

gender gender of child (1 = male, 0 = female)

birthord birth order of child

magebirth Age of mother at birth of child

bthwht whether child was born low birth weight (1 = yes, 0 = no)

breastfed whether child was breastfed (1 = yes, 0 = no)

medu Highest grade completed by child's mother

math PIAT Math Standard Score

read PIAT Reading Recognition Standard Score

hhnum Number of household members in household

select_group_tbl	<i>Summarize multiple response variables by group</i>
------------------	---

Description

select_group_tbl() presents frequency counts and percentages (count, percent) for binary (e.g., Unselected/Selected) and ordinal (e.g., strongly disagree to strongly agree) variables with the same variable stem by some grouping variable. A variable stem is a common prefix found in related variable names, often corresponding to similar survey items, that represents a shared concept before unique identifiers (like time points) are added. For example, in the stem_social_psych dataset, the two variables belong_belongStem_w1 and belong_belongStem_w2 share the variable stem belong_belongStem (e.g., "I feel like an outsider in STEM"), with suffixes (_w1, _w2) indicating different measurement waves. By default, missing data are excluded from the calculations in a listwise fashion.

Usage

```
select_group_tbl(
  data,
  var_stem,
  group,
  escape_stem = FALSE,
  ignore_stem_case = FALSE,
  group_type = "variable",
  group_name = NULL,
  escape_group = FALSE,
  ignore_group_case = FALSE,
  remove_group_non_alnum = TRUE,
  na_removal = "listwise",
  pivot = "longer",
  only = NULL,
  var_labels = NULL,
  ignore = NULL
)
```

Arguments

data	A data frame.
var_stem	A character string of a variable stem or the full name of a variable in data.
group	A character string of a variable in data or a pattern to use to search for variables in data.
escape_stem	A logical value indicating whether to escape var_stem. Default is FALSE.
ignore_stem_case	A logical value indicating whether the search for columns matching the supplied var_stem is case-insensitive. Default is FALSE.

<code>group_type</code>	A character string that defines the type of grouping variable. Should be one of <code>pattern</code> or <code>variable</code> . Default is <code>variable</code> , in which case the variable matching the group string will be searched for within data.
<code>group_name</code>	A character string piped to the final table to replace the name of group.
<code>escape_group</code>	A logical value indicating whether to escape string supplied to group.
<code>ignore_group_case</code>	A logical value indicating whether group is case- insensitive. Default is <code>FALSE</code> .
<code>remove_group_non_alnum</code>	A logical value indicating whether to remove all non-alphanumeric characters (anything that is not a letter or number) from group. Default is <code>TRUE</code> .
<code>na_removal</code>	A character string specifying how to remove missing values. Should be one of <code>pairwise</code> or <code>listwise</code> . Default is <code>listwise</code> .
<code>pivot</code>	A character string specifying the format of the returned summary table. The default is <code>longer</code> , which returns the data in long format. To return the data in wide format, use <code>wider</code> .
<code>only</code>	A character string or vector of character strings of the kinds of summary data to return. Default is <code>NULL</code> , which returns counts (<code>count</code>) and percentages (<code>percent</code>).
<code>var_labels</code>	An optional named character vector or list where each element maps labels to variable names. If any element is unnamed or if any labels do not match variables in returned from data, all labels will be ignored and the table will be printed without them.
<code>ignore</code>	An optional named vector or list specifying values to exclude from the dataset and analysis. By default, <code>NULL</code> includes all available values. To omit values from variables returned by <code>var_stem</code> , use the provided stem as the name. To exclude values from both <code>var_stem</code> variables and a grouping variable in data, supply a list.

Value

A tibble displaying frequency counts and/or percentages for each value of a set of variables sharing the same variable stem. The results are grouped by either a grouping variable in the data or by a pattern matched with variable names. When the output is in the wider format, columns beginning with `count_value` and `percent_value` prefixes report the count and percentage, respectively, for each distinct response value of the variable within each group.

Author(s)

Ama Nyame-Mensah

Examples

```
select_group_tbl(data = stem_social_psych,
  var_stem = "belong_belong",
  group = "\\d",
  group_type = "pattern",
  group_name = "wave",
  na_removal = "pairwise",
```

```

        pivot = "wider",
        only = "count")

tas_recoded <-
  tas |>
  dplyr::mutate(sex = dplyr::case_when(
    sex == 1 ~ "female",
    sex == 2 ~ "male",
    TRUE ~ NA)) |>
  dplyr::mutate(dplyr::across(
    .cols = dplyr::starts_with("involved_"),
    .fns = ~ dplyr::case_when(
      .x == 1 ~ "selected",
      .x == 0 ~ "unselected",
      TRUE ~ NA)
  ))

select_group_tbl(data = tas_recoded,
  var_stem = "involved_",
  group = "sex",
  group_type = "variable",
  na_removal = "pairwise",
  pivot = "wider")

depressive_recoded <-
  depressive |>
  dplyr::mutate(sex = dplyr::case_when(
    sex == 1 ~ "male",
    sex == 2 ~ "female",
    TRUE ~ NA)) |>
  dplyr::mutate(dplyr::across(
    .cols = dplyr::starts_with("dep_"),
    .fns = ~ dplyr::case_when(
      .x == 1 ~ "often",
      .x == 2 ~ "sometimes",
      .x == 3 ~ "hardly",
      TRUE ~ NA
    )
  ))

select_group_tbl(data = depressive_recoded,
  var_stem = "dep",
  group = "sex",
  group_type = "variable",
  na_removal = "listwise",
  pivot = "wider",
  only = "percent",
  var_labels =
    c("dep_1" = "how often child feels sad and blue",
      "dep_2" = "how often child feels nervous, tense, or on edge",
      "dep_3" = "how often child feels happy",
      "dep_4" = "how often child feels bored",
      "dep_5" = "how often child feels lonely",

```

```
"dep_6" = "how often child feels tired or worn out",
"dep_7" = "how often child feels excited about something",
"dep_8" = "how often child feels too busy to get everything"))
```

select_tbl

Summarize multiple response variables

Description

select_tbl() presents frequency counts and percentages (count, percent) for binary (e.g., Unselected/Selected) and ordinal (e.g., strongly disagree to strongly agree) variables with the same variable stem. A variable stem is a common prefix found in related variable names, often corresponding to similar survey items, that represents a shared concept before unique identifiers (like time points) are added. For example, in the stem_social_psych dataset, the two variables belong_belongStem_w1 and belong_belongStem_w2 share the variable stem belong_belongStem (e.g., "I feel like an outsider in STEM"), with suffixes (_w1, _w2) indicating different measurement waves. By default, missing data are excluded from the calculations in a listwise fashion.

Usage

```
select_tbl(
  data,
  var_stem,
  escape_stem = FALSE,
  ignore_stem_case = FALSE,
  na_removal = "listwise",
  pivot = "longer",
  only = NULL,
  var_labels = NULL,
  ignore = NULL
)
```

Arguments

data	A data frame.
var_stem	A character string of a variable stem or the full name of a variable in data.
escape_stem	A logical value indicating whether to escape var_stem. Default is FALSE.
ignore_stem_case	A logical value indicating whether the search for columns matching the supplied var_stem is case-insensitive. Default is FALSE.
na_removal	A character string specifying how to remove missing values. Should be one of pairwise or listwise. Default is listwise.
pivot	A character string specifying the format of the returned summary table. The default is longer, which returns the data in long format. To return the data in wide format, use wider.

only	A character string or vector of character strings of the kinds of summary data to return. Default is NULL, which returns counts (count) and percentages (percent).
var_labels	An optional named character vector or list where each element maps labels to variable names. If any element is unnamed or if any labels do not match variables in returned from data, all labels will be ignored and the table will be printed without them.
ignore	An optional vector that contains values to exclude from the data. Default is NULL, which includes all present values.

Value

A tibble displaying frequency counts and/or percentages for each value of a set of variables sharing the same variable stem. When the output is in the wider format, columns beginning with count_value and percent_value prefixes report the count and percentage, respectively, for each distinct response value of the variable.

Author(s)

Ama Nyame-Mensah

Examples

```
select_tbl(data = tas,
           var_stem = "involved_",
           na_removal = "pairwise")

select_tbl(data = depressive,
           var_stem = "dep",
           na_removal = "listwise",
           pivot = "wider",
           only = "percent")

var_label_example <-
  c("dep_1" = "how often child feels sad and blue",
    "dep_2" = "how often child feels nervous, tense, or on edge",
    "dep_3" = "how often child feels happy",
    "dep_4" = "how often child feels bored",
    "dep_5" = "how often child feels lonely",
    "dep_6" = "how often child feels tired or worn out",
    "dep_7" = "how often child feels excited about something",
    "dep_8" = "how often child feels too busy to get everything")

select_tbl(data = depressive,
           var_stem = "dep",
           na_removal = "pairwise",
           pivot = "longer",
           var_labels = var_label_example)

select_tbl(data = depressive,
           var_stem = "dep",
           na_removal = "pairwise",
```

```

pivot = "wider",
only = "count",
var_labels = var_label_example)

```

social_psy_data

Social Psychological (Generated) Data

Description

These data were generated to produce social psychological data applicable to real-world contexts.

Usage

```
social_psy_data
```

Format

A data.frame with 10,200 rows and 17 columns:

id participant id number)

belong_1 I feel like I belong at this institution (1=Strongly Disagree, 2=Disagree,3=Neither agree nor disagree,4=Agree,5=Strongly Agree)

belong_2 I feel like part of the community (1=Strongly Disagree, 2=Disagree,3=Neither agree nor disagree,4=Agree,5=Strongly Agree)

belong_3 I feel valued by this institution (1=Strongly Disagree, 2=Disagree,3=Neither agree nor disagree,4=Agree,5=Strongly Agree)

identity_1 This institution is a big part of who I am (1=Strongly Disagree,2=Disagree,3=Neither agree nor disagree,4=Agree,5=Strongly Agree)

identity_2 I feel comfortable being myself in this setting (1=Strongly Disagree,2=Disagree,3=Neither agree nor disagree,4=Agree, 5=Strongly Agree)

identity_3 This institution is a big part of who I am (1=Strongly Disagree, 2=Disagree,3=Neither agree nor disagree,4=Agree,5=Strongly Agree)

identity_4 I care about doing well at this institution (1=Strongly Disagree, 2=Disagree,3=Neither agree nor disagree,4=Agree,5=Strongly Agree)

selfEfficacy_1 I am confident about A (1=Strongly Disagree,2=Disagree, 3=Neither agree nor disagree,4=Agree,5=Strongly Agree)

selfEfficacy_2 I am confident about B (1=Strongly Disagree,2=Disagree, 3=Neither agree nor disagree,4=Agree,5=Strongly Agree)

selfEfficacy_3 I am confident about C (1=Strongly Disagree,2=Disagree, 3=Neither agree nor disagree,4=Agree,5=Strongly Agree)

selfEfficacy_4 I am confident about D (1=Strongly Disagree,2=Disagree, 3=Neither agree nor disagree,4=Agree,5=Strongly Agree)

selfEfficacy_5 I am confident about E (1=Strongly Disagree,2=Disagree, 3=Neither agree nor disagree,4=Agree,5=Strongly Agree)

selfEfficacy_6 I am confident about F (1=Strongly Disagree,2=Disagree, 3=Neither agree nor disagree,4=Agree,5=Strongly Agree)

selfEfficacy_7 I am confident about G (1=Strongly Disagree,2=Disagree, 3=Neither agree nor disagree,4=Agree,5=Strongly Agree)

gender Participant's gender identity (1=Woman,2=Man,3=Non-binary, 4=Self-identify,5=Transgender,6=Gender-queer/non-conforming)

citizen Participant's citizenship status (1=U.S. citizen,2=Non-U.S. citizen with permanent residency,3=Non-U.S. citizen with temporary visa,4=Other)

stem_social_psych *STEM Social Psychological (Generated) Data*

Description

These data were generated to produce social psychological data applicable to a subset of college students participating in a Science, Technology, Engineering, and Mathematics (STEM) intervention program.

Usage

stem_social_psych

Format

A data.frame with 786 rows and 37 columns:

id student id number)

belong_belongStem_w1 I feel like I belong in STEM (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree,4=Somewhat agree,5=Strongly agree)

belong_outsiderStem_w1 I feel like an outsider in STEM (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree,4=Somewhat agree,5=Strongly agree)

identity_identityStem_w1 STEM is a big part of who I am. (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree,4=Somewhat agree,5=Strongly agree)

belong_welcomedStem_w1 I feel welcomed in STEM workplaces (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree,4=Somewhat agree,5=Strongly agree)

identity_noCommonStem_w1 I do not have much in common with the other students in my STEM classes.(1=Strongly disagree,2=Somewhat disagree,3=Neither disagree nor agree, 4=Somewhat agree, 5=Strongly agree)

selfEfficacy_passStemCourses_w1 pass my STEM courses.(1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree,4=Somewhat agree, 5=Strongly agree)

selfEfficacy_learnConcepts_w1 learn the foundations and concepts of scientific thinking. (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree, 4=Somewhat agree, 5=Strongly agree)

- selfEfficacy_stemField_w1** do well in a stem-related field. (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree,4=Somewhat agree,5=Strongly agree)
- selfEfficacy_learnScience_w1** quickly learn new science areas, systems, techniques or concepts on my own. (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree, 4=Somewhat agree, 5=Strongly agree)
- selfEfficacy_contributeProject_w1** contribute to a science project. (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree,4=Somewhat agree, 5=Strongly agree)
- selfEfficacy_commScience_w1** clearly communicate scientific problems and findings to varied audiences (1=Strongly disagree,2=Somewhat disagree, 3=Neither disagree nor agree, 4=Somewhat agree,5=Strongly agree)
- selfEfficacy_scientist_w1** become a scientist. (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree,4=Somewhat agree,5=Strongly agree)
- selfEfficacy_completeUG_w1** complete an undergraduate STEM degree. (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree,4=Somewhat agree, 5=Strongly agree)
- selfEfficacy_admitGrad_w1** get admitted to a graduate STEM program. (1=Strongly disagree,2=Somewhat disagree,3=Neither disagree nor agree,4=Somewhat agree, 5=Strongly agree)
- selfEfficacy_successGrad_w1** be successful in a graduate STEM program. (1=Strongly disagree,2=Somewhat disagree,3=Neither disagree nor agree,4=Somewhat agree, 5=Strongly agree)
- belong_belongStem_w2** I feel like I belong in STEM (1=Strongly disagree, 2=Somewhat disagree, 3=Neither disagree nor agree,4=Somewhat agree,5=Strongly agree)
- belong_outsiderStem_w2** I feel like an outsider in STEM. (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree,4=Somewhat agree,5=Strongly agree)
- identity_identityStem_w2** STEM is a big part of who I am. (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree,4=Somewhat agree,5=Strongly agree)
- belong_welcomedStem_w2** I feel welcomed in STEM workplaces. (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree,4=Somewhat agree,5=Strongly agree)
- identity_noCommonStem_w2** I do not have much in common with the other students in my STEM classes.(1=Strongly disagree,2=Somewhat disagree,3=Neither disagree nor agree, 4=Somewhat agree, 5=Strongly agree)
- selfEfficacy_passStemCourses_w2** pass my STEM courses. (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree,4=Somewhat agree, 5=Strongly agree)
- selfEfficacy_learnConcepts_w2** learn the foundations and concepts of scientific thinking. (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree, 4=Somewhat agree, 5=Strongly agree)
- selfEfficacy_stemField_w2** do well in a stem-related field. (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree,4=Somewhat agree,5=Strongly agree)
- selfEfficacy_learnScience_w2** quickly learn new science areas, systems, techniques or concepts on my own. (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree, 4=Somewhat agree, 5=Strongly agree)
- selfEfficacy_contributeProject_w2** contribute to a science project. (1=Strongly disagree, 2=Somewhat disagree,3=Neither disagree nor agree,4=Somewhat agree, 5=Strongly agree)
- selfEfficacy_commScience_w2** clearly communicate scientific problems and findings to varied audiences (1=Strongly disagree,2=Somewhat disagree, 3=Neither disagree nor agree, 4=Somewhat agree,5=Strongly agree)

- selfEfficacy_scientist_w2** become a scientist. (1=Strongly disagree, 2=Somewhat disagree, 3=Neither disagree nor agree, 4=Somewhat agree, 5=Strongly agree)
- selfEfficacy_completeUG_w2** complete an undergraduate STEM degree. (1=Strongly disagree, 2=Somewhat disagree, 3=Neither disagree nor agree, 4=Somewhat agree, 5=Strongly agree)
- selfEfficacy_admitGrad_w2** get admitted to a graduate STEM program. (1=Strongly disagree, 2=Somewhat disagree, 3=Neither disagree nor agree, 4=Somewhat agree, 5=Strongly agree)
- selfEfficacy_successGrad_w2** be successful in a graduate STEM program. (1=Strongly disagree, 2=Somewhat disagree, 3=Neither disagree nor agree, 4=Somewhat agree, 5=Strongly agree)
- is_male** Participant's current sex (0=Not Male, 1=Male)
- has_disability** Whether participant has a disability (0=No, 1=Yes)
- firstGen** Whether participant is a first generation college student (0=No, 1=Yes)
- stemMajor** Whether participant is a STEM Major (0=No, 1=Yes)
- expLearning** Whether student has participated in an experiential learning program, such as an internship, research, or leadership opportunity. (0=No, 1=Yes)
- urm** Whether participant is Asian, Middle Eastern/Arab or White (0) vs. Black, Indigenous, Hispanic/Latino, or Mixed Race (1)

tas

*Panel Study of Income Dynamics (PSID) Transition into Adulthood
Supplement (TAS) Data*

Description

These data are a subset from the Panel Study of Income Dynamics (PSID) Transition into Adulthood Supplement. The data contains 2,526 observations and 8 variables.

For more information about the Panel Study of Income Dynamics, visit <https://psidonline.isr.umich.edu/CDS/default.aspx>.

Usage

tas

Format

A tibble with 2,526 rows and 8 columns:

pid personal identification number)

sex sex of individual (1 = female, 2 = male)

involved_arts whether the individual participated in any organized activities related to art, music, or the theater in the last 12 months (1 = yes, 0 = no)

involved_sports whether the individual was a member of any athletic or sports teams in the last 12 months (1 = yes, 0 = no)

involved_schoolClubs whether the individual was involved with any high school or college clubs or student government in the last 12 months (1 = yes, 0 = no)

involved_election whether the individual voted in the national election in November 2016 that was held to elect the President (1 = yes, 0 = no)

involved_socialActionGrps whether the individual was involved in any political groups, solidarity or ethnic-support groups or social-action groups in the last 12 months (1 = yes, 0 = no)

involved_volunteer whether the individual was involved in any unpaid volunteer or community service work in the last 12 months (1 = yes, 0 = no)

Index

* datasets

- depressive, 5
- nlsy, 10
- social_psy_data, 16
- stem_social_psych, 17
- tas, 19

- cat_group_tbl, 2
- cat_tbl, 3
- check_named_vctr, 4

- depressive, 5

- mean_group_tbl, 6
- mean_tbl, 8

- nlsy, 10

- select_group_tbl, 11
- select_tbl, 14
- social_psy_data, 16
- stem_social_psych, 17

- tas, 19