

Package ‘yowie’

September 6, 2021

Title Longitudinal Wages Data from the National Longitudinal Survey of Youth 1979

Version 0.1.0

Description Longitudinal wages data sets and several demographic variables from the National Longitudinal Survey of Youth from 1979 to 2018. There are three data sets in this package: The wages data from the cohort whose highest grade completed is up to high school; The wages data of the high school dropouts and; The demographic data of the cohort in the survey year 1979.

License GPL (>= 3)

Encoding UTF-8

LazyData true

RoxygenNote 7.1.1

URL <https://github.com/numbats/yowie>

BugReports <https://github.com/numbats/yowie/issues>

Depends R (>= 3.5.0), tsibble (>= 1.0.1)

Suggests ggplot2, knitr, markdown, rmarkdown, dplyr, tidyr, forcats, janitor, kableExtra

VignetteBuilder knitr

NeedsCompilation no

Author Dewi Amaliah [aut, cre],
Di Cook [aut],
Emi Tanaka [aut]

Maintainer Dewi Amaliah <dlamaleeah@gmail.com>

Repository CRAN

Date/Publication 2021-09-06 09:20:02 UTC

R topics documented:

demog_nlsy79	2
wages	2
wages_hs_do	4

Index**6**

demog_nlsy79	<i>Demographic Data from the National Longitudinal Survey of Youth (NLSY79)</i>
--------------	---

Description

A data set contains demographic information of the NLSY79 cohort. The cohort included 12,686 respondents ages 14-22 when first interviewed in 1979.

Format

A data frame contains 12,686 rows and 7 variables:

id A unique individual's ID number.

age_1979 The age of the subject in 1979.

gender Gender of the subject, FEMALE and MALE.

race Race of the subject, NON-BLACK, NON-HISPANIC; HISPANIC; BLACK.

hgc Highest grade completed.

hgc_i Integer of highest grade completed.

yr_hgc The year when the highest grade completed.

Source

The U.S. Bureau of Labor Statistics. (2021, January 6). *National Longitudinal Survey of Youth 1979*. <https://www.nlsinfo.org/content/cohorts/nlsy79/get-data>

wages	<i>Wages Data from the National Longitudinal Survey of Youth (NLSY79)</i>
-------	---

Description

A data set contains longitudinal data of mean hourly wages along with several demographic variables of Americans from the National Longitudinal Survey of Youth (NLSY79) held by the U.S. Bureau of Labor Statistics from Round 1 (1979 survey year) to Round 28 (2018 survey year). The cohort provided in this data set is the highest grade completed, up to 12th grade, and participated in at least five rounds of surveys.

Format

A tibble with 103,994 rows and 15 variables:

id A unique individual's ID number. This is the key of the data.

year The year the observation was taken. This could be the index of the data.

wage The mean of the hourly wages the individual gets at each of their different jobs. The value could be a weighted or an arithmetic mean. The weighted mean is used when the information of hours of work as the weight is available. The mean hourly wage could also be a predicted value if the original value is considered influential by the robust linear regression as part of data cleaning.

age_1979 The age of the subject in 1979.

gender Gender of the subject, FEMALE and MALE.

race Race of the subject, NON-BLACK, NON-HISPANIC; HISPANIC; BLACK.

hgc Highest grade completed.

hgc_i Integer of highest grade completed.

yr_hgc The year when the highest grade completed.

njobs Number of jobs that an individual has.

hours The total number of hours the individual usually works per week.

is_wm Whether the mean hourly wage is weighted mean, using the hour work as the weight, or regular/arithmetic mean. TRUE = is weighted mean. FALSE = is regular mean.

is_pred Whether the mean hourly wage is a predicted value or not.

Source

The U.S. Bureau of Labor Statistics. (2021, January 6). *National Longitudinal Survey of Youth 1979*. <https://www.nlsinfo.org/content/cohorts/nlsy79/get-data>

Examples

```
# data summary
wages

library(ggplot2)
library(dplyr)
library(tibble)
wages_ids <- key_data(wages) %>% select(id)
wages %>%
  dplyr::filter(id %in% sample_n(wages_ids, 10)$id) %>%
  ggplot() +
    geom_line(aes(x = year,
                  y = wage,
                  group = id), alpha = 0.8)
```

wages_hs_do

Wages Data of High School Dropout from the National Longitudinal Survey of Youth (NLSY79)

Description

A data set contains longitudinal data of mean hourly wages along with several demographic variables from the National Longitudinal Survey of Youth (NLSY79) held by the U.S. Bureau of Labor Statistics from Round 1 (1979 survey year) to Round 28 (2018 survey year). The cohort provided in this data set is high school dropouts, i.e., the male aged between 14 and 17 years and only completed either 9th, 10th, 11th grade, or aged at least 19 years old when completed high school and participated in at least five rounds of surveys.

Format

A tibble contains 97,087 rows and 15 variables:

#'

id A unique individual's ID number. This is the key of the data.

year The year the observation was taken. This could be the index of the data.

wage The mean of the hourly wages the individual gets at each of their different jobs. The value could be a weighted or an arithmetic mean. The weighted mean is used when the information of hours of work as the weight is available. The mean hourly wage could also be a predicted value if the original value is considered influential by the robust linear regression as part of data cleaning.

age_1979 The age of the subject in 1979.

gender Gender of the subject, FEMALE and MALE.

race Race of the subject, NON-BLACK, NON-HISPANIC; HISPANIC; BLACK.

hgc Highest grade completed.

hgc_i Integer of highest grade completed.

yr_hgc The year when the highest grade completed.

njobs Number of jobs that an individual has.

hours The total number of hours the individual usually works per week.

is_wm Whether the mean hourly wage is weighted mean, using the hour work as the weight, or regular/arithmetic mean. TRUE = is weighted mean. FALSE = is regular mean.

is_pred Whether the mean hourly wage is a predicted value or not.

Source

The U.S. Bureau of Labor Statistics. (2021, January 6). *National Longitudinal Survey of Youth 1979*. <https://www.nlsinfo.org/content/cohorts/nlsy79/get-data>

wages_hs_do

5

Examples

```
# saving the data into a new object
library(yowie)
library(tsibble)
wages_hs_do %>% key_data()
```

Index

demog_nlsy79, 2

wages, 2

wages_hs_do, 4