Package: learningtower (via r-universe)

Title OECD PISA Datasets from 2000-2018 in an Easy-to-Use Format

Version 1.0.1

Description The Programme for International Student Assessment (PISA) is a global study conducted by the Organization for Economic Cooperation and Development (OECD) in member and non-member countries to assess educational systems by assessing 15-year-old school students academic performance in mathematics, science, and reading. This datasets contains information on their scores and other socioeconomic characteristics, information about their school and its infrastructure, as well as the countries that are taking part in the program.

Depends R (>= 3.5.0)

Encoding UTF-8

RoxygenNote 7.2.3

VignetteBuilder knitr

License MIT + file LICENSE

URL https://kevinwang09.github.io/learningtower/,
https://github.com/kevinwang09/learningtower

BugReports https://github.com/kevinwang09/learningtower/issues

Imports tibble, dplyr

Suggests testthat (>= 3.0.0), knitr, rmarkdown, ggplot2, forcats, scales

Config/testthat/edition 3

Repository https://kevinwang09.r-universe.dev

RemoteUrl https://github.com/kevinwang09/learningtower

RemoteRef HEAD

RemoteSha c8d7af020a771f99570d6bc4a77451c1f0a592ea

Contents

countrycode	*Country iso3c and name mapping for PISA OECD countries participants.*

Description

A dataset containing mapping of the country ISO code to the country names. More information on participating countries can be found at https://www.oecd.org/pisa/aboutpisa/pisa-participants.htm.

Format

A tibble of the following variables

- country: Country 3 character code. Note that some regions/territories are coded as country for ease of input. Character.

- country_name: Country name. Note that some regions/territories are coded as country for ease of input. Character.

load_student	*load_student() function allows the user to extract the PISA student scores for any desired year from 2000-2018*

Description

load_student() function was created to extract the data of student's scores in any years from 2000-2018, the function requires any of the year as it argument or a string "all" that will return all the PISA scores of the students from the years 2000-2018.

Usage

```
load_student(year = "2000")
```

Arguments

year   is the required parameter for the function to display the dataset the user wants to view the PISA scores for the selected year else the entire student data will be available to the user

Value

A dataset of PISA scores of students that took the test in the selected year as per user from the years 2000-2018

Examples

```
## Not run:
library(learningtower)
student_all <- load_student("all")
student_2000 <- load_student("2000")

## End(Not run)
```

| school | *School data available for the years 2000-2018 from the PISA OECD database* |

Description

A dataset containing school weight and other information from the triennial testing of 15 year olds around the globe. Original data available from https://www.oecd.org/pisa/data/.

Format

A tibble of the following variables

- year: Year of the PISA data. Factor.
- country_iso3c: Country 3 character code. Note that some regions/territories are coded as country for ease of input. Factor.
- school_id: The school identification number, unique for each country and year combination. Factor.
- fund_gov: Percentage of total funding for school year from government. Numeric.
- fund_fees: Percentage of total funding for school year from student fees or school charges paid by parents. Numeric.
- fund_donation: Percentage of total funding for school year from benefactors, donations, bequests, sponsorship, parent fundraising. Numeric.
- enrol_boys: Number of boys in the school. Numeric.
- enrol_girls: Number of girls in the school. Numeric.
- stratio: Student-Teacher ratio. Numeric.
- public_private: Is the school a public or private school. Factor.
- staff_shortage: Shortage of staff. Numeric.
- sch_wgt: The final survey weight score for the schools. Numeric.
- school_size: The school size. Numeric.

student_subset_2000	*Sample student data available for the years 2000-2018 from the PISA OECD database*

Description

A sample student subset dataset containing scores and other information from the triennial testing of 15 year olds around the globe. Original data available from <https://www.oecd.org/pisa/data/>.

Format

A tibble of the following variables

- year: Year of the PISA data. Factor.
- country: Country 3 character code. Note that some regions/territories are coded as country for ease of input. Factor.
- school_id: The school identification number, unique for each country and year combination. Factor.
- student_id: The student identification number, unique for each school, country and year combination. Factor.
- mother_educ: Highest level of mother's education. Ranges from "less than ISCED1" to "ISCED 3A". Factor. Note that in 2000, all entries are missing.
- father_educ: Highest level of father's education. Ranges from "less than ISCED1" to "ISCED 3A". Factor. Note that in 2000, all entries are missing.
- gender: Gender of the student. Only "male" and "female" are recorded. Factor. Note that we call this variable gender and not sex as this term was used in the OECD PISA database.
- computer: Possession of computer. Only "yes" and "no" are recorded. Factor.
- internet: Access to internet. Only "yes" and "no" are recorded. Factor.
- math: Simulated score in mathematics. Numeric.
- read: Simulated score in reading. Numeric.
- science: Simulated score in science. Numeric.
- stu_wgt: The final survey weight score for the student score. Numeric.
- desk: Possession of desk to study at. Only "yes" and "no" are recorded. Factor.
- room: Possession of a room of your own. Only "yes" and "no" are recorded. Factor.
- dishwasher: Possession of a dishwasher. Only "yes" and "no" are recorded. Factor. Note that in 2015 and 2018, all entries are missing.
- television: Number of televisions. "0", "1", "2" are code for no, one and two TVs in the house. "3+" codes for three or more TVs. Factor. Note that in 2003, all entries are missing.
- computer_n: Number of computers. "0", "1", "2" are code for no, one and two computers in the house. "3+" codes for three or more computers. Factor. Note that in 2003, all entries are missing.

- car: Number of cars. "0", "1", "2" are code for no, one and two cars in the house. "3+" codes for three or more cars Factor. Note that in 2003, all entries are missing.

- book: Number of books. Factor. Note that encoding is different in the years 2000 and 2003 compared to all other years. Factor. Evaluate `table(student$book, student$year)` for a demo.

- wealth: Family wealth. Numeric. Note that in 2003, all entries are missing.

- escs: Index of economic, social and cultural status. Numeric.

Examples

```
library(dplyr)
data(student_subset_2000)
data(student_subset_2003)
dplyr::bind_rows(
student_subset_2000,
student_subset_2003
)
```

Index