Package 'collegeScorecard'

July 22, 2025

Title US College Scorecard Data

Version 0.2.0

Description Provides a tidied subset of the US College Scorecard dataset,
containing institutional characteristics, enrollment, student aid,
costs, and student outcomes at institutions of higher education in the
United States.

License CC0

URL <https://github.com/gadenbuie/scorecard-db>,
<https://collegescorecard.ed.gov/data/>

BugReports <https://github.com/gadenbuie/scorecard-db/issues>

Depends R (>= 3.5)

Config/testthat/edition 3

Encoding UTF-8

LazyData true

RoxygenNote 7.3.2

NeedsCompilation no

Author Garrick Aden-Buie [aut, cre, cph] (ORCID:
<https://orcid.org/0000-0002-7111-0077>)

Maintainer Garrick Aden-Buie <garrick@adenbuie.com>

Repository CRAN

Date/Publication 2025-02-13 14:20:06 UTC

Contents

| school | *College Scorecard: School Data* |

Description

Information about institutions in the U.S. College Scorecard dataset. The uses the most recent report for each institution. The school dataset can be used to filter the scorecard dataset. Join the two datasets using the id column.

The data set contain only colleges that participate in Title IV federal financial aid programs. In the tidy data set, column names have been changed for readability and consistency.

You can find the original data set at the U.S. Department of Education's College Scorecard website. The code for the data transformation process is available at `https://github.com/gadenbuie/scorecard-db`.

Usage

    school

Format

A data frame with 11,300 rows and 25 variables. Original column names from the source dataset are noted in parenthesis.

id [integer] A unique identifier for each institution. (UNITID)

name [character] Institution name. (INSTNM)

city [character] City. (CITY)

state [character] State postcode. (STABBR)

zip [character] ZIP code. (ZIP)

latitude [numeric] Latitude. (LATITUDE)

longitude [numeric] Longitude. (LONGITUDE)

url [character] URL for institution's homepage. (INSTURL)

deg_predominant [factor] Predominant undergraduate degree awarded. (PREDDEG)

deg_highest [factor] Highest degree awarded. (HIGHDEG)

control [factor] Control of institution (IPEDS). (CONTROL)

locale_type [factor] Locale of institution. (LOCALE)

locale_size [factor] Locale of institution. (LOCALE)

adm_req_test [factor] Test score requirements for admission. (ADMCON7)

is_hbcu [logical] Flag for Historically Black College and University. (HBCU)

is_pbi [logical] Flag for predominantly black institution. (PBI)

is_annhi [logical] Flag for Alaska Native Native Hawaiian serving institution. (ANNHI)

is_tribal [logical] Flag for tribal college and university. (TRIBAL)

is_aanapii [logical] Flag for Asian American Native American Pacific Islander-serving institution. (AANAPII)

is_hsi [logical] Flag for Hispanic-serving institution. (HSI)

`is_nanti [logical]` Flag for Native American non-tribal institution. (`NANTI`)

`is_only_men [logical]` Flag for men-only college. (`MENONLY`)

`is_only_women [logical]` Flag for women-only college. (`WOMENONLY`)

`is_only_distance [logical]` Flag for distance-education-only education. (`DISTANCEONLY`)

`religious_affiliation [factor]` Religious affiliation of the institution. (`RELAFFIL`)

References

https://collegescorecard.ed.gov/data/

---

scorecard   *College Scorecard: Scorecard Data*

Description

Historical data from the U.S. College Scorecard dataset. The data includes historical information on the number of students, average cost, median earnings after graduations, rates of admission and completion, and ACT and SAT test scores.

The data set contain only colleges that participate in Title IV federal financial aid programs. In the tidy data set, column names have been changed for readability and consistency.

You can find the original data set at the U.S. Department of Education's College Scorecard website. The code for the data transformation process is available at https://github.com/gadenbuie/scorecard-db.

Usage

```
scorecard
```

Format

A data frame with 183,306 rows and 24 variables. Original column names from the source dataset are noted in parenthesis.

`id [integer]` A unique identifier for each institution. (`UNITID`)

`academic_year [character]` The academic year of the record.

`n_undergrads [integer]` Enrollment of undergraduate certificate/degree-seeking students. (`UGDS`)

`cost_tuition_in [numeric]` In-state tuition and fees. (`TUITIONFEE_IN`)

`cost_tuition_out [numeric]` Out-of-state tuition and fees. (`TUITIONFEE_OUT`)

`cost_books [numeric]` Cost of attendance: estimated books and supplies. (`BOOKSUPPLY`)

`cost_room_board_on [numeric]` Cost of attendance: on-campus room and board. (`ROOMBOARD_ON`)

`cost_room_board_off [numeric]` Cost of attendance: off-campus room and board. (`ROOMBOARD_OFF`)

`cost_avg [numeric]` Average net price for Title IV institutions (private for-profit and nonprofit institutions). (`NPT4_PRIV`)

`cost_avg_income_0_30k [numeric]` Average net price for $0-$30,000 family income (private for-profit and nonprofit institutions). (`NPT41_PRIV`)

cost_avg_income_30_48k [numeric] Average net price for $30,001-$48,000 family income (private for-profit and nonprofit institutions). (NPT42_PRIV)

cost_avg_income_48_75k [numeric] Average net price for $48,001-$75,000 family income (private for-profit and nonprofit institutions). (NPT43_PRIV)

cost_avg_income_75_110k [numeric] Average net price for $75,001-$110,000 family income (private for-profit and nonprofit institutions). (NPT44_PRIV)

cost_avg_income_110k_plus [numeric] Average net price for $110,000+ family income (private for-profit and nonprofit institutions). (NPT45_PRIV)

amnt_earnings_med_10y [numeric] Median earnings of students working and not enrolled 10 years after entry. (MD_EARN_WNE_P10)

rate_completion [numeric] Completion rate for first-time, full-time students at four-year institutions (100% of expected time to completion). (C100_4)

rate_admissions [numeric] Admission rate for all campuses rolled up to the 6-digit OPE ID. (ADM_RATE_ALL)

score_sat_avg [numeric] Average SAT equivalent score of students admitted. (SAT_AVG)

score_act_p25 [numeric] 25th percentile of the ACT cumulative score. (ACTCM25)

score_act_p75 [numeric] 75th percentile of the ACT cumulative score. (ACTCM75)

score_sat_verbal_p25 [numeric] 25th percentile of SAT scores at the institution (critical reading). (SATVR25)

score_sat_verbal_p75 [numeric] 75th percentile of SAT scores at the institution (critical reading). (SATVR75)

score_sat_math_p25 [numeric] 25th percentile of SAT scores at the institution (math). (SATMT25)

score_sat_math_p75 [numeric] 75th percentile of SAT scores at the institution (math). (SATMT75)

References

<https://collegescorecard.ed.gov/data/>

Index