

Package: collegeScorecard (via r-universe)

October 16, 2026

Title US College Scorecard Data

Version 0.3.0.9000

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/roxygen2/version 8.1.1

Config/testthat/edition 3

Encoding UTF-8

LazyData true

LazyDataCompression xz

Roxygen list(markdown = TRUE)

Repository <https://gadenbuie.r-universe.dev>

Date/Publication 2026-10-16 18:24:01 UTC

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

RemoteRef main

RemoteSha 250aeaf02b551d8dc72dedb40f7157f689822c20

RemoteSubdir pkg

Contents

school	2
scorecard	3

Index	6
--------------	----------

 school

College Scorecard: School Data

Description

Information about institutions in the U.S. College Scorecard dataset. This 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.

Schools are included only if their most recent record has OPEFLAG = 1 (participates in Title IV federal financial aid programs). Branch campuses with OPEFLAG = 2, deferment-only institutions, and nonparticipating institutions are excluded. 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](https://nces.ed.gov/ipeds/data/college-scorecard/). The code for the data transformation process is available at <https://github.com/gadenbuie/scorecard-db>.

Usage

```
school
```

Format

A data frame with 11,465 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, costs of attendance and average net prices, median earnings after entry, rates of admission and completion, and ACT and SAT test scores.

Schools are included only if their most recent record has OPEFLAG = 1 (participates in Title IV federal financial aid programs). Branch campuses with OPEFLAG = 2, deferment-only institutions, and nonparticipating institutions are excluded. Column names have been changed for readability and consistency.

cost_avg and cost_avg_income_* are average net prices, not costs of attendance. The cost_attendance_* columns report costs of attendance for academic-year and program-year institutions separately. rate_completion is the four-year institution completion rate within 100% of normal time; rate_completion_150_4yr and rate_completion_150_1t4yr report completion within 150% of normal time for four-year and less-than-four-year institutions.

The Most Recent Institution bulk file can draw different metrics from different academic years. A single academic_year in this dataset need not match every metric in that file.

You can find the original data set at the U.S. Department of Education's [College Scorecard website](https://collegescorecard.ed.gov/). The code for the data transformation process is available at <https://github.com/gadenbuie/scorecard-db>.

Usage

scorecard

Format

A data frame with 200,699 rows and 28 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_attendance_academic_year [numeric] Average cost of attendance (academic year institutions). (COSTT4_A)
 cost_attendance_program_year [numeric] Average cost of attendance (program-year institutions). (COSTT4_P)
 cost_avg [numeric] Average net price for Title IV institutions (public, private nonprofit, and private for-profit). Uses the private value when available, otherwise the public value. (NPT4_PRIV, NPT4_PUB)
 cost_avg_income_0_30k [numeric] Average net price for \$0-\$30,000 family income (public and private institutions). Uses the private value when available, otherwise the public value. (NPT41_PRIV, NPT41_PUB)
 cost_avg_income_30_48k [numeric] Average net price for \$30,001-\$48,000 family income (public and private institutions). Uses the private value when available, otherwise the public value. (NPT42_PRIV, NPT42_PUB)
 cost_avg_income_48_75k [numeric] Average net price for \$48,001-\$75,000 family income (public and private institutions). Uses the private value when available, otherwise the public value. (NPT43_PRIV, NPT43_PUB)
 cost_avg_income_75_110k [numeric] Average net price for \$75,001-\$110,000 family income (public and private institutions). Uses the private value when available, otherwise the public value. (NPT44_PRIV, NPT44_PUB)
 cost_avg_income_110k_plus [numeric] Average net price for \$110,000+ family income (public and private institutions). Uses the private value when available, otherwise the public value. (NPT45_PRIV, NPT45_PUB)
 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_completion_150_4yr [numeric] Completion rate for first-time, full-time students at four-year institutions (150% of expected time to completion). (C150_4)
 rate_completion_150_lt4yr [numeric] Completion rate for first-time, full-time students at less-than-four-year institutions (150% of expected time to completion). (C150_L4)
 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

* **datasets**

school, [2](#)

scorecard, [3](#)

school, [2](#)

scorecard, [3](#)