Package: edfinr (via r-universe)

October 20, 2026

Type Package

Title Access Tidy Education Finance Data

Version 0.2.0

Description Provides easy access to tidy education finance data using Bellwether's methodology to combine NCES F-33 Survey, Census Bureau Small Area Income Poverty Estimates (SAIPE), community data from the ACS 5-Year Estimates, the Comparable Wage Index for Teachers (CWIFT) from NCES EDGE, and school district land area from U.S. Census Bureau Gazetteer Files. The package simplifies downloading, caching, and filtering education finance data by year and state, enabling researchers and analysts to explore K-12 education funding patterns, revenue sources, expenditure categories, and demographic factors across U.S. school districts.

License CC BY 4.0

URL https://github.com/bellwetherorg/edfinr, https://bellwetherorg.github.io/edfinr/

BugReports https://github.com/bellwetherorg/edfinr/issues

Depends R (>= 4.1.0)

Imports cli, dplyr, nanoparquet (>= 0.5.0), rlang, tibble, utils

Encoding UTF-8

Roxygen list(markdown = TRUE)

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

VignetteBuilder knitr

Config/testthat/edition 3

Config/Needs/website sf, tigris

Config/roxygen2/version 8.0.0

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

Date/Publication 2026-08-21 17:52:54 UTC

RemoteUrl https://github.com/bellwetherorg/edfinr

RemoteRef HEAD

RemoteSha 7c7b8340effbc05ad7f56421054066ae93b50046

Contents

get_finance_data 2
get_states 4
list_variables 4

Index **6**

`get_finance_data` *Get Education Finance Data*

Description

This function downloads tidy education finance data using data from the NCES F-33 Survey, Census Bureau Small Area Income Poverty Estimates (SAIPE), and community data from the ACS 5-Year Estimates.

Usage

```
get_finance_data(
  yr = "2023",
  geo = "all",
  dataset_type = "skinny",
  cpi_adj = "none",
  refresh = FALSE,
  quiet = FALSE
)
```

Arguments

`yr`	A string specifying the year(s) to retrieve. Can be a single year ("2023"), a range ("2020:2023"), or "all" for all available years. Defaults to "2023". Only the requested year(s) are downloaded – each year is a separate hosted file of roughly 3-6 MB, so a single-year request transfers far less than the full panel. `yr = "all"` fetches the entire history from one combined file. When `cpi_adj` names a year outside the request, that year's file is also downloaded to source the baseline, then dropped from the returned data.
`geo`	A string specifying the geographic scope. Can be "all" for all states (default), a single state code ("KY"), or a comma-separated list of state codes ("IN,KY,OH,TN").
`dataset_type`	A string specifying whether to download the "skinny" (default) or "full" dataset. The skinny version excludes detailed expenditure data for faster downloads.

`cpi_adj` A string specifying the CPI adjustment baseline year. Can be "none" (default) for no adjustment, or a year between 2012-2023 to use as the baseline year. When a year is specified (e.g., "2023"), revenue, expenditure, and economic variables are adjusted to that school year's dollars using CPI averaged over the months of the school year (e.g., "2023" uses the 2022-23 school year CPI). Capital outlay and debt-interest flows are adjusted; debt and fund-balance stocks (`debt_*`, `fund_bal_*`) and the CWIFT index are returned nominal. When cpi_adj is set to a value other than "none", a new column "cpi_adj_index" will be added to the output showing the adjustment index used for each row.

`refresh` A logical value indicating whether to force a refresh of the cached data. Default is FALSE.

`quiet` A logical value indicating whether to suppress download progress messages. Default is FALSE.

Details

Downloaded files are cached for the duration of the R session in a subdirectory of `tempdir()`, so repeated calls in one session do not re-download; the cache is cleared when the session ends. Use `refresh = TRUE` to force a fresh download. During downloads the package temporarily raises R's download timeout to at least 600 seconds (the `yr = "all"` combined files are 38-54 MB); a higher user-set `options(timeout = )` is respected.

Value

A tibble containing the requested education finance data.

Examples

```
# Check valid parameters without downloading
get_states()  # Valid state codes


# These examples require internet access and may take time to download

# get data for Kentucky for 2023
ky_data <- get_finance_data(yr = "2023", geo = "KY")

# get data for multiple years
ky_multi <- get_finance_data(yr = "2021:2023", geo = "KY")

# get full dataset with detailed expenditure data
ky_full <- get_finance_data(yr = "2023", geo = "KY", dataset_type = "full")

# get data adjusted to 2023 dollars
ky_adjusted <- get_finance_data(yr = "2021:2023", geo = "KY", cpi_adj = "2023")

# get data for multiple states for several years
regional_data <- get_finance_data(yr = "2021:2023", geo = "IN,KY,OH,TN")
```

get_states *Get list of valid state codes*

Description

Returns the valid two-letter state codes that can be used with get_finance_data

Usage

```
get_states()
```

Value

A character vector of state codes

Examples

```
# Get all valid state codes
states <- get_states()
head(states)
```

list_variables *List available variables in the education finance dataset*

Description

This function provides information about the variables available in the education finance dataset, including their names, types, and brief descriptions.

Usage

```
list_variables(dataset_type = "skinny", category = "all")
```

Arguments

dataset_type	A string specifying whether to list variables for "skinny" (default) or "full" dataset.
category	Optional. Filter variables by category: "id", "time", "geographic", "demographic", "revenue", "expenditure", "economic", "governance", "debt", "cwift", or "all" (default).

Value

A tibble with one row per variable and the columns `name`, `type`, `category`, `source`, `f33_item`, `first_yr_avail`, and `description`. `f33_item` gives the F-33 survey item(s) a variable is drawn from (e.g., "F12", "TCAPOUT / V33"); it is `NA` for variables from non-F-33 sources and for edfinr-adjusted measures (such as the adjusted revenue variables), whose construction is described in the "Data Sources and Methodology" vignette.

Examples

```
# list all available variables in skinny dataset
vars <- list_variables()
head(vars)

# list all variables in full dataset
full_vars <- list_variables(dataset_type = "full")
nrow(full_vars)

# list only expenditure variables in full dataset
exp_vars <- list_variables(dataset_type = "full", category = "expenditure")
head(exp_vars)
```

Index

get_finance_data, 2
get_states, 4

list_variables, 4

tempdir(), *3*