Package 'FFdownload'

May 7, 2026

Type Package

Title Download Data from Kenneth French's Website

Version 1.2.0

Description Downloads all the datasets (you can exclude the daily ones or specify a list of those you are targeting specifically) from Kenneth French's Website at <https://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html>, process them and convert them to list of 'xts' (time series).

Depends R (>= 3.5.0), utils, stats, rvest, xts, xml2, zoo

Imports timetk

License MIT + file LICENSE

URL https://github.com/sstoeckl/ffdownload, https://sstoeckl.github.io/ffdownload/

BugReports https://github.com/sstoeckl/ffdownload/issues

Encoding UTF-8

RoxygenNote 7.3.3

Suggests knitr, rmarkdown, dplyr, viridis, ggplot2, tidyr, tibble, purrr, testthat (>= 3.0.0)

VignetteBuilder knitr

Config/testthat/edition 3

NeedsCompilation no

Author Sebastian Stoeckl [aut, cre] (ORCID: <https://orcid.org/0000-0002-4196-6093>, Package commissioner and maintainer.), Annar Massimov [ctb] (Original developer of FFdownload.)

Maintainer Sebastian Stoeckl <sebastian.stoeckl@uni.li>

Repository CRAN

Date/Publication 2026-02-25 22:50:02 UTC

Contents

`FFdownload` *Downloads Datasets from Kenneth French's Website*

Description

`FFdownload` returns an RData file with all (possibility to exclude the large daily) datasets from Kenneth French's Website. Should help researchers to work with the datasets and update the regularly. Allows for reproducible research. Be aware that processing (especially when including daily files) takes quite a long time!

Usage

```
FFdownload(
  output_file = "data.Rdata",
  tempd = NULL,
  exclude_daily = FALSE,
  download = TRUE,
  download_only = FALSE,
  listsave = NULL,
  inputlist = NULL,
  format = "xts",
  na_values = NULL,
  return_data = FALSE,
  action = NULL,
  cache_days = Inf,
  match_threshold = 0.3
)
```

Arguments

`output_file`	name of the .RData file to be saved (include path if necessary)
`tempd`	specify if you want to keep downloaded files somewhere save. Seems to be necessary for reproducible research as the files on the website do change from time to time
`exclude_daily`	excludes the daily datasets (are not downloaded) ==> speeds the process up considerably
`download`	set to TRUE if you actually want to download again. set to false and specify tempd to keep processing the already downloaded files

download_only — set to FALSE if you want to process all your downloaded files at once

listsave — if not NULL, the list of unzipped files is saved here (good for processing only a limited number of files through inputlist). Is written before inputlist is processed.

inputlist — if not NULL, FFdownload tries to match the names from the list with the list of zip-files

format — (set to xts) specify "xts" or "tbl"/"tibble" for the output format of the nested lists

na_values — numeric vector of sentinel values to replace with NA (e.g. c(-99, -999, -99.99)). French's files use -99.99 or -999 to denote missing observations. Default NULL preserves the original behaviour (no replacement).

return_data — logical. If TRUE, the FFdata list is returned invisibly in addition to being saved to output_file. Default FALSE preserves the original behaviour.

action — convenience alternative to the download/download_only flag pair. One of "all" (download + process), "list_only" (save file list only), "download_only" (download but do not process), or "process_only" (process already-downloaded files from tempd). When action is provided it overrides download and download_only. Default NULL retains the original flag-based behaviour.

cache_days — numeric. When greater than 0 and less than Inf, zip files already present in tempd that are younger than cache_days days are reused instead of being re-downloaded. Default Inf preserves the original behaviour (never re-download an existing file).

match_threshold — numeric in [0,1]. If the similarity between a requested inputlist entry and its fuzzy-matched filename is below this threshold a warning is emitted. Use FFmatch() to inspect matches before downloading. Default 0.3.

Value

Invisibly returns the FFdata list when return_data = TRUE; otherwise called for its side-effect of writing an RData file.

Examples

```
## Not run:
tempf <- tempfile(fileext = ".RData"); outd <- paste0(tempdir(),"/",format(Sys.time(), "%F_%H-%M"))
temptxt <- tempfile(fileext = ".txt")

# Example 1: Use FFdownload to get a list of all monthly zip-files. Save that list as temptxt.

FFdownload(exclude_daily=TRUE,download=FALSE,download_only=TRUE,listsave=temptxt)
read.delim(temptxt,sep = ",")
# set vector with only files to download (we try a fuzzyjoin, so "Momentum" should be enough to get
# the Momentum Factor)
inputlist <- c("Research_Data_Factors","Momentum_Factor","ST_Reversal_Factor","LT_Reversal_Factor")
# Now process only these files if they can be matched (download only)
FFdownload(exclude_daily=FALSE,tempd=outd,download=TRUE,download_only=FALSE,
inputlist=inputlist,output_file = tempf)
list.files(outd)
```

```
# Then process all the downloaded files
FFdownload(output_file = tempf, exclude_daily=TRUE,tempd=outd,download=FALSE,
download_only=FALSE,inputlist=inputlist)
load(tempf); FFdata$`x_F-F_Momentum_Factor`$monthly$Temp2[1:10]

# Example 2: Use action parameter and return data directly

FFdata <- FFdownload(
  inputlist = c("F-F_Research_Data_5_Factors_2x3"),
  output_file = tempf,
  action = "all",
  na_values = c(-99, -999, -99.99),
  return_data = TRUE
)
FFdata$`x_F-F_Research_Data_5_Factors_2x3`$monthly$Temp2

## End(Not run)
```

FFget *Download and return a single French dataset directly*

Description

FFget is a convenience wrapper around FFdownload that downloads one named dataset and returns it directly into the R session — no intermediate .RData file, no load() call required.

The function uses all of FFdownload's parsing engine, so every sub-table present in the original CSV (value-weighted returns, equal-weighted returns, number of firms, etc.) is available in the returned list.

Usage

```
FFget(
  name,
  frequency = "monthly",
  subtable = NULL,
  exclude_daily = TRUE,
  na_values = c(-99, -999, -99.99),
  format = "tbl"
)
```

Arguments

name	character. The dataset name as it appears in FFlist()$name, e.g. "F-F_Research_Data_Factors" or "F-F_Momentum_Factor". Fuzzy matching is applied, so partial names work (check with FFmatch first).
frequency	character. Which frequency sub-list to extract. One of "monthly" (default), "annual", or "daily". Set to NULL to return all three frequencies as a named list.

subtable	character. Name of the sub-table within the chosen frequency, e.g. "Temp2" or "annual_factors:_january-december". Set to NULL (default) to return all sub-tables as a named list.
exclude_daily	logical. Passed to FFdownload. Default TRUE.
na_values	numeric vector of sentinel values to replace with NA. Defaults to c(-99, -999, -99.99) — the values French uses for missing observations. Set to NULL to disable replacement.
format	character. "tbl" (default) or "xts".

Value

A tibble, xts object, or named list, depending on frequency, subtable, and format.

Examples

```
## Not run:
# Get the main monthly Fama-French 3-factor table directly as a tibble
ff3 <- FFget("F-F_Research_Data_Factors", subtable = "Temp2")
head(ff3)

# Get all sub-tables for the 5-factor model
ff5_all <- FFget("F-F_Research_Data_5_Factors_2x3", subtable = NULL)
names(ff5_all)

# Get annual data as xts
ff3_ann <- FFget("F-F_Research_Data_Factors", frequency = "annual", format = "xts")

## End(Not run)
```

FFlist *List available datasets on Kenneth French's website*

Description

FFlist scrapes Kenneth French's data library and returns a data frame (or tibble) of available datasets with their names and download URLs. This replaces the listsave workaround in FFdownload and makes the dataset inventory directly usable with dplyr::filter() or View().

Usage

```
FFlist(exclude_daily = TRUE)
```

Arguments

exclude_daily	logical. If TRUE (default), daily datasets are excluded from the returned list.

Value

A data frame (or tibble if the **tibble** package is available) with columns:

name Dataset name, as used in `inputlist` and as key in the `FFdata` list (without the leading `x_` prefix and without the `_CSV.zip` suffix).

file_url Full HTTPS URL of the zip file.

is_daily Logical flag indicating whether the dataset contains daily data. Only present when `exclude_daily = FALSE`.

Examples

```
## Not run:
# Browse all available monthly/annual datasets
fl <- FFlist()
head(fl, 10)

# Include daily datasets
FFlist(exclude_daily = FALSE)

# Filter with dplyr
library(dplyr)
FFlist() |> filter(grepl("Momentum", name))

## End(Not run)
```

FFmatch *Preview fuzzy-matching results before downloading*

Description

`FFmatch` shows how each entry in `inputlist` would be matched to an available dataset by the fuzzy-matching logic inside `FFdownload`. Use this to verify matches before triggering a download, especially when dataset names are abbreviated or partially specified.

Usage

```
FFmatch(inputlist, exclude_daily = TRUE)
```

Arguments

`inputlist`	character vector of (partial) dataset names to match, as you would pass to the `inputlist` argument of `FFdownload`.
`exclude_daily`	logical. If `TRUE` (default), daily datasets are excluded from the candidate pool.

Value

A data frame (or tibble) with one row per entry in `inputlist` and columns:

requested The input string as supplied.

matched The dataset name that would be selected by `FFdownload`.

edit_distance Raw Levenshtein edit distance between `requested` and `matched`.

similarity 1 - edit_distance / nchar(matched), clamped to [0, 1]. Values below 0.3 suggest a potentially wrong match.

Examples

```
## Not run:
FFmatch(c("Research_Data_Factors", "Momentum", "ST_Reversal"))

## End(Not run)
```

Index