

Package ‘ropercenter’

July 2, 2023

Type Package

Title Reproducible Data Retrieval from the Roper Center Data Archive

Version 0.3.2

Date 2023-06-29

URL <https://github.com/fsolt/ropercenter>,
<https://fsolt.org/ropercenter/>

BugReports <https://github.com/fsolt/ropercenter/issues>

Description Reproducible, programmatic retrieval of datasets from the Roper Center data archive. The Roper Center for Public Opinion Research <<https://ropercenter.cornell.edu>> maintains the largest archive of public opinion data in existence, but researchers using these datasets are caught in a bind. The Center's terms and conditions bar redistribution of downloaded datasets, but to ensure that one's work can be reproduced, assessed, and built upon by others, one must provide access to the raw data one employed. The `ropercenter` package cuts this knot by providing registered users with programmatic, reproducible access to Roper Center datasets from within R.

License MIT + file LICENSE

Suggests knitr, rmarkdown

VignetteBuilder knitr

Imports RSelenium, dplyr, foreign, haven, magrittr, netstat, purrr,
readr, rio, stringr, tibble, tidyr

Encoding UTF-8

RoxygenNote 7.2.3

NeedsCompilation no

Author Frederick Solt [aut, cre, cph],
Jennifer Lin [ctb],
Paul Gronke [ctb],
Dave Peterson [ctb]

Maintainer Frederick Solt <frederick-solt@uiowa.edu>

Repository CRAN

Date/Publication 2023-07-02 17:00:02 UTC

R topics documented:

read_ascii	2
roper_download	4

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

read_ascii	<i>Read ASCII datasets downloaded from the Roper Center</i>
------------	---

Description

read_ascii helps format ASCII data files downloaded from the Roper Center.

Usage

```
read_ascii(
  file,
  total_cards = 1,
  var_names,
  var_cards = 1,
  var_positions,
  var_widths,
  card_pattern,
  respondent_pattern
)
```

Arguments

file	A path to an ASCII data file.
total_cards	For multicard files, the number of cards in the file.
var_names	A string vector of variable names.
var_cards	For multicard files, a numeric vector of the cards on which var_names are recorded.
var_positions	A numeric vector of the column positions in which var_names are recorded.
var_widths	A numeric vector of the widths used to record var_names.
card_pattern	For use when the file does not contain a line for every card for every respondent (or contains extra lines that correspond to no respondent), a regular expression that matches the file's card identifier; e.g., if the card number is stored in the last digit of each line, "\d\$".
respondent_pattern	For use when the file does not contain a line for every card for every respondent (or contains extra lines that correspond to no respondent), a regular expression that matches the file's respondent identifier; e.g., if the respondent number is stored in the first four digits of each line, preceded by a space, "(?<=^\\s)d4".

Details

Many older Roper Center datasets are available only in ASCII format, which is notoriously difficult to work with. The `read_ascii` function facilitates the process of extracting selected variables from ASCII datasets. For single-card files, one can simply identify the names, positions, and widths of the needed variables from the codebook and pass them to `read_ascii`'s `var_names`, `var_positions`, and `var_widths` arguments. Multicard datasets are more complicated. In the best case, the file contains one line per card per respondent; then, the user can extract the needed variables by adding only the `var_cards` and `total_cards` arguments. When this condition is violated—there is not a line for every card for every respondent, or there are extra lines—the function will throw an error and request the user specify the additional arguments `card_pattern` and `respondent_pattern`.

Value

A data frame containing any variables specified in the `var_names` argument, plus a numeric respondent identifier and as many string card variables (`card1`, `card2`, ...) as specified by the `total_cards` argument.

Examples

```
## Not run:
# a single-card file
roper_download("USAIP01982-1197G", # Gallup Poll for June 25-28, 1982
              download_dir = tempdir()) # remember to specify a directory for your download

gallup1982 <- read_ascii(file = file.path(tempdir(), "USAIP01982-1197G",
                                         "1197.dat"),
                      var_names = c("q09j", "weight"),
                      var_positions = c(38, 1),
                      var_widths = c(1, 1))

# a multi-card file, with extra lines that make the card_pattern and
# respondent_pattern arguments necessary
roper_download("USAIPOCNUS1996-9603008", # Gallup/CNN/USA Today Poll: Politics/1996 Election
              download_dir = tempdir()) # remember to specify a directory for your download

gallup1996 <- read_ascii(file = file.path(tempdir(), "USAIPOCNUS1996-9603008",
                                         "a9603008.dat"),
                      var_names = c("q43a", "q44", "weight"),
                      var_cards = c(6, 6, 1),
                      var_positions = c(62, 64, 13),
                      var_widths = c(1, 1, 3),
                      total_cards = 7,
                      card_pattern = "(?<=^.{10})\\d",
                                # (a digit, preceded by the start of the line
                                # and ten other characters)
                      respondent_pattern = "(?<=^\\s{2})\\d{4}"
                                # (# four digits, preceded by the start of the line
                                # and two whitespace characters)

## End(Not run)
```

roper_download	<i>Download datasets from the Roper Center</i>
----------------	--

Description

roper_download provides a programmatic and reproducible means to download datasets from the Roper Center's data archive

Usage

```
roper_download(
  file_id,
  affiliation = getOption("roper_affiliation"),
  email = getOption("roper_email"),
  password = getOption("roper_password"),
  reset = FALSE,
  download_dir = "roper_data",
  msg = TRUE,
  convert = TRUE,
  delay = 2
)
```

Arguments

file_id	The unique identifier (or optionally a vector of these identifiers) for the dataset(s) to be downloaded (see details). Both new Archive Numbers and old Historical Archive Numbers, if the latter are available, may be used.
affiliation, email, password	Your Roper Center affiliation, email, and password (see details)
reset	If TRUE, you will be asked to re-enter your Roper Center affiliation, email, and password.
download_dir	The directory (relative to your working directory) to which files from the Roper Center will be downloaded.
msg	If TRUE, outputs a message showing which data set is being downloaded.
convert	If TRUE, converts downloaded file(s) to .RData format.
delay	If the speed of your connection to the Roper Center data archive is particularly slow, roper_download may encounter problems. Increasing the delay parameter may help.

Details

To avoid requiring others to edit your scripts to insert their own affiliation, email, and password or to force them to do so interactively, the default is set to fetch this information from the user's .Rprofile. Before running roper_download, then, you should be sure to add these options to your .Rprofile (use `usethis::edit_r_profile()` is one particularly easy way), substituting your own info for the example below:

```
options("roper_affiliation" = "Upper Midwest University", "roper_email" = "juanita-herrara@uppermidwes  
"roper_password" = "password123!")
```

Value

The function returns nothing, but has the side effect of downloading all files of the datasets identified in the `file_id` argument.

Examples

```
## Not run:  
roper_download(file_id = c("31117412", "USPEW2015-GOVERNANCE"),  
              download_dir = tempdir()) # remember to specify a directory for your download  
  
## End(Not run)
```

Index

`read_ascii`, 2
`roper_download`, 4