Package 'CEOdata'

January 20, 2025

Title Datasets of the CEO (Centre d'Estudis d'Opinio) for Opinion
Polls in Catalonia

Version 1.3.1.1

Description Easy and convenient access to the datasets of the
``Centre d'Estudis d'Opinio'', the Catalan institution for polling and public
opinion. The package uses the data stored in the servers of the CEO and returns
it in a tidy format.

Depends R (>= 4.1.0), haven, dplyr (>= 1.0.0), urltools, stringr,
jsonlite

Suggests knitr, R.rsp, ggplot2, tidyr (> 1.1.1), curl, rmarkdown,
vtable, compareGroups

License GPL-3

URL https://ceo.gencat.cat/ca/inici/,
https://github.com/ceopinio/CEOdata/

BugReports https://github.com/ceopinio/CEOdata/issues/

Encoding UTF-8

RoxygenNote 7.2.3

VignetteBuilder knitr, R.rsp

NeedsCompilation no

Author Xavier Fernández-i-Marín [aut, cre]
(<https://orcid.org/0000-0002-9522-8870>)

Maintainer Xavier Fernández-i-Marín <xavier.fim@gmail.com>

Repository CRAN

Date/Publication 2023-07-06 14:00:12 UTC

Contents

| CEOdata | *Import datasets / microdata from the "Centre d'Estudis d'Opinio"* |

Description

Easy and convenient access to the datasets / microdata of the "Centre d'Estudis d'Opinio", the Catalan institution for polling and public opinion. The package uses the data stored in the servers of the CEO and returns it in a tidy format (tibble).

Easy and convenient access to the datasets / microdata of the "Centre d'Estudis d'Opinio", the catalan institution for polling and public opinion. The package uses the data stored in the servers of the CEO and returns it in a tidy format (tibble).

Usage

```
CEOdata(
  kind = "barometer",
  reo = NA,
  raw = FALSE,
  extra_variables = TRUE,
  date_start = NA,
  date_end = NA
)
```

Arguments

kind   Character vector with the sort of microdata required. Defaults to "barometer",
                     that contains the whole set of Barometers from 2014 (presential interviews).
                     "barometer_until_2013" contains the interviews performed by phone until 2013,
                     with a somewhat different questionnaire and structure. For such dataset you
                     need a third-party software installed in your computer to be able to uncompress
                     the RAR original file. It is the option by default. But if a specific reo study is
                     requested in the `reo` argument, then the `kind` argument does not apply anymore
                     and only a specific study is retrieved.

reo   Character vector of length one that allows to get the dataset of a specific REO
                     study (Registre d'Estudis d'Opinio, the internal register ID used by the CEO) to
                     download. By default (when `reo = NA`) it uses the `kind` argument. Not all the
                     studies carried on by the CEO (and therefore listed in the `CEOmeta()` function
                     call) have microdata available. Only the ones that return TRUE to the column
                     `microdata_available` in `CEOmeta()`.

raw   Logical value to indicate if SPSS labels are transformed into factors. Defaults to
                     FALSE. Otherwise, when TRUE, it returns the matrices as imported by haven::read_spss()
                     without modification. Does not apply to data from singular REOs, only to
                     barometers retrieved using `kind`.

extra_variables
                     Logical value as to whether include (default) complementary variables such as
                     date (Data). Defaults to TRUE. Names of such new variables only use upper

case in the first letter. Extra variables are added at the end. Does not apply to data from singular REOs, only to barometers retrieved using `kind`.

date_start   Character vector with a starting date ("YYYY-MM-DD") for the data. It only applies to the barometers retrieved using `kind`, not to other studies.

date_end   Character vector with an end date ("YYYY-MM-DD") for the data. It only applies to the barometers retrieved using `kind`, not to other studies.

Details

It works either by specifying the kind of merged barometer (using the `kind` argument), or either providing a singular study (using the `reo` argument).

Value

A tibble with the individuals' responses to the questionnaire retrieved.

References

<http://xavier-fim.net/packages/CEOdata/>.

Examples

```
## Not run:
# Get the merged barometer from 2014, by default (assume kind = "barometer").
d <- CEOdata()

# Get the number of individuals surveyed and the number of variables recorded.
dim(d)

# Get the identifiers of the different Barometers retrieved
unique(d$BOP_NUM)

# Get a specific study
d746 <- CEOdata(reo = "746")

## End(Not run)
```

CEOmeta   *Import metadata from the "Centre d'Estudis d'Opinio"*

Description

Easy and convenient access to the metadata of the "Centre d'Estudis d'Opinio", the Catalan institution for polling and public opinion. It allows to search for specific terms to obtain the details of the datasets available

Usage

```
CEOmeta(
  reo = NULL,
  search = NULL,
  date_start = NA,
  date_end = NA,
  browse = FALSE,
  browse_translate = NULL,
  browse_force = FALSE
)
```

Arguments

reo	Character vector of length one that allows to get the metadata only of a specific REO (Registre d'Estudis d'Opinio, the internal register ID used by the CEO) to download. When not NULL it has precedence with the search, date_start and date_end arguments.
search	Character vector with keywords to look for within several columns of the CEO metadata (title, summary, objectives and tags -descriptors-). Each element of the vector is strictly evaluated (all words are considered to be found in the format they appear, like in "AND"), while by using several elements in the vector the search works like an "OR" clause. Lower or upper cases are not considered.
date_start	Character vector with a starting date ("YYYY-MM-DD") for the data.
date_end	Character vector with an end date ("YYYY-MM-DD") for the data.
browse	Logical value. When turned to TRUE, the browser opens the URLs of the required surveys. Only a maximum of 10 entries are opened.
browse_translate	
	When opening the relevant entries in the browser (browse must be TRUE), use automatic translation to the language specified using Google Translate ('oc' for Occitan/Aranese, 'de' to German, 'en' to English, 'eu' to Basque, 'gl' for Galician or 'sp' to Spanish).
browse_force	Logical value. When TRUE it overcomes the limitation of only opening a maximum of 10 URLs. Use it with caution.

Value

A tibble with the metadata of the surveys produced by the CEO.

Examples

```
## Not run:
# Retrieve the metadata of the surveys ever produced by the CEO:
meta <- CEOmeta()
dim(meta)

# Search for specific terms in any of the metadata fields
# in this case, "internet".
CEOmeta(search = "internet")
```

```
# now for the combination of "Medi" AND "Ambient"
CEOmeta(search = "Medi ambient")

# now for the combination of ("Medi" AND "Ambient") OR "Municipi"
CEOmeta(search = c("Medi ambient", "Municipi"))

# Search for all registers starting in 2020
CEOmeta(date_start = "2020-01-01")

# Get the entry for a specific study (REO) and open its description in a browser
CEOmeta(reo = "746", browse = TRUE)

## End(Not run)
```

CEOsearch	*Search for keywords in the labels of variables and responses of the survey data*

Description

Easy and convenient access to the metadata of the "Centre d'Estudis d'Opinio", the Catalan institution for polling and public opinion. It allows to search for specific terms to obtain the details of the datasets available

Usage

```
CEOsearch(d, keyword = NULL, where = "variables", translate = FALSE)
```

Arguments

d	Microdata retrieved from the CEO using the CEOdata() function. It is a data frame with variable labels.
keyword	The character string defining the word / concept to look for within the microdata.
where	A character vector specifying if the function should look amongst variable labels ("variables", default), or amongst value labels ("values").
translate	Logical. When TRUE, it opens a browser with an automatic translation to English of the variable names and labels using Google Translate. Given the specificity of the terms, only the English translation is provided. Defaults to FALSE.

Value

A tibble with the set of variables that match the keyword ("Variable"). If the variables are requested, the second variable is their labels ("Label"), and if the values are required the second on is the value labels ("Value").

Examples

```
## Not run:
# Retrieve a dataset to use the function
d <- CEOdata()

# Get the whole set of variable labels
CEOsearch(d)

# Get the whole set of value labels
CEOsearch(d, where = "values")

# Search for specific variable names and variable labels with the string "edat" (age).
CEOsearch(d, keyword = "edat")

# Search for specific variable names and variable labels with the string "edat" (age),
# and translate the results to English.
CEOsearch(d, keyword = "edat", translate = TRUE)

# now for the combination of "valoracio" OR "covid" OR "govern".
CEOsearch(d, keyword = c("valoració", "covid",  "govern"))

## End(Not run)
```

Index