

Package ‘eurostat’

March 27, 2015

Type Package

Title Tools for Eurostat Open Data

Version 1.0.16

Date 2015-03-27

MailingList rOpenGov <ropengov-forum@googlegroups.com>

Description Tools to download data from the Eurostat database
<http://ec.europa.eu/eurostat/> together with search and manipulation utilities.

License BSD_2_clause + file LICENSE

Depends R (>= 3.0.2)

Imports plyr, tidyr

Suggests knitr, mapproj, plotrix, reshape2, rmarkdown, ggplot2,
testthat

URL <https://github.com/ropengov/eurostat>

BugReports <https://github.com/ropengov/eurostat/issues>

VignetteBuilder knitr

NeedsCompilation no

Repository CRAN

Author Lahti Leo [aut, cre],
Biecek Przemyslaw [aut],
Kainu Markus [aut],
Huovari Janne [aut]

Maintainer Lahti Leo <louhos@googlegroups.com>

Date/Publication 2015-03-27 14:16:46

R topics documented:

eurostat-package	2
clean_eurostat_cache	3

eurotime2date	3
eurotime2num	4
get_eurostat	5
get_eurostat_dic	6
get_eurostat_raw	7
get_eurostat_toc	8
label_eurostat	9
search_eurostat	10

Index	12
--------------	-----------

eurostat-package	<i>R Tools for Eurostat open data</i>
------------------	---------------------------------------

Description

Brief summary of the eurostat package

Details

Package:	eurostat
Type:	Package
Version:	See sessionInfo() or DESCRIPTION file
Date:	2014-2015
License:	BSD_2_clause + LICENSE
LazyLoad:	yes

R Tools for Eurostat Open Data

Author(s)

Leo Lahti, Przemyslaw Biecek, Janne Huovari, Markus Kainu <ropengov-forum@googlegroups.com>
<http://github.com/ropengov/eurostat>

References

See citation("eurostat") <http://github.com/ropengov/eurostat>

Examples

```
library(eurostat)
```

clean_eurostat_cache *Clean Eurostat Cache*

Description

Delete all .rds files from the eurostat cache directory. See [get_eurostat](#) for more on cache.

Usage

```
clean_eurostat_cache(cache_dir = NULL)
```

Arguments

cache_dir A path to cache directory. If NULL (default) tries to clean default temporary cache directory.

Examples

```
clean_eurostat_cache()
```

eurotime2date *Date conversion from Eurostat time format*

Description

A function to convert Eurostat time values to objects of class [Date](#) representing calendar dates.

Usage

```
eurotime2date(x, last = FALSE)
```

Arguments

x a character string with time information in Eurostat time format.
last a logical. If FALSE (default) the date is the first date of the period (month, quarter or year). If TRUE the date is the last date of the period.

Value

an object of class [Date](#).

Author(s)

Janne Huovari <ropengov-forum@googlegroups.com> <http://github.com/ropengov/eurostat>

Examples

```
## Not run:
lp <- get_eurostat("namq_aux_lp", time_format = "raw")
lp$time <- eurotime2date(x = lp$time)

un <- get_eurostat("une_rt_m", time_format = "raw")
un$time <- eurotime2date(x = un$time)

lpa <- get_eurostat("nama_aux_lp", time_format = "raw")
lpa$time <- eurotime2date(x = lpa$time)

## End(Not run)
```

eurotime2num

A conversion of a Eurostat time format to numeric

Description

A function to convert Eurostat time values to a numeric. Bi-annual, quarterly and monthly data is presented as fraction of the year in beginning of the period.

Usage

```
eurotime2num(x)
```

Arguments

x a character string with time information in Eurostat time format.

Value

see [as.numeric](#).

Author(s)

Janne Huovari <ropengov-forum@googlegroups.com> <http://github.com/ropengov/eurostat>

Examples

```
## Not run:
lp <- get_eurostat("namq_aux_lp", time_format = "raw")
lp$time <- eurotime2num(x = lp$time)

un <- get_eurostat("une_rt_m", time_format = "raw")
un$time <- eurotime2num(x = un$time)

lpa <- get_eurostat("nama_aux_lp", time_format = "raw")
lpa$time <- eurotime2num(x = lpa$time)

## End(Not run)
```

get_eurostat	<i>Read data from Eurostat database.</i>
--------------	--

Description

Download dataset from the Eurostat database (ec.europa.eu/eurostat).

Usage

```
get_eurostat(id, time_format = "date", cache = TRUE, update_cache = FALSE,
             cache_dir = NULL)
```

Arguments

id	A code name for the data set of interest. See the table of contents of eurostat datasets for details.
time_format	a string giving a type of the conversion of the time column from the eurostat format. A "date" (default) converts to a Date with a first date of the period. A "date_last" converts to a Date with a last date of the period. A "num" converts to a numeric and "raw" does not do conversion. See eurotime2date and eurotime2num .
cache	a logical wheather to do caching. Default is TRUE.
update_cache	a locigal wheater to update cache. Can be set also with options(eurostat_update = TRUE)
cache_dir	a path to cache directory. The NULL uses directory eurostat directory in the temporary directory from tempdir . Directory can be set also with option eurostat_cache_dir.

Details

Datasets are downloaded from the Eurostat bulk download facility <http://ec.europa.eu/eurostat/estat-navtree-portlet-prod/BulkDownloadListing>. The data is transformed into the molten row-column-value format (RCV).

By default datasets are cached. In a temporary directory by default or in a named directory if cache_dir or option eurostat_cache_dir is defined.

Value

a data.frame. One column for each dimension in the data and the values column for numerical values. The time column for a time dimension.

See Also

[{tidy_eurostat}](#)

Examples

```
## Not run:
k <- get_eurostat("namq_aux_lp")
k <- get_eurostat("namq_aux_lp", update_cache = TRUE)
dir.create("r_cache")
k <- get_eurostat("namq_aux_lp", cache_dir = "r_cache")
options(eurostat_update = TRUE)
k <- get_eurostat("namq_aux_lp")
options(eurostat_cache_dir = "r_cache")
k <- get_eurostat("namq_aux_lp")
k <- get_eurostat("namq_aux_lp", cache = FALSE)

## End(Not run)
```

get_eurostat_dic

Download a Eurostat dictionary

Description

Download a dictionary for given coded variable from Eurostat (ec.europa.eu/eurostat). Dictionaries link codes with human-readable labels. To translate codes to labels with [label_eurostat](#).

Arguments:

Usage

```
get_eurostat_dic(dictname)
```

```
getEurostatDictionary(dictname)
```

Arguments

dictname Character, dictionary for given variable name will be downloaded.

Returns:

Value

A data.frame with two columns, first with code names and second with full names.

Functions

- `getEurostatDictionary`: Old deprecated version

Author(s)

Przemyslaw Biecek and Leo Lahti <ropengov-forum@googlegroups.com>

References

see citation("eurostat").

See Also

[label_eurostat](#), [get_eurostat](#), [search_eurostat](#).

Examples

```
## Not run:
  tmp <- get_eurostat_dic("crop_pro")
  head(tmp)

## End(Not run)
```

get_eurostat_raw

Download data from the eurostat database

Description

Downloads datasets from the eurostat database and transforms into tabular format.

Usage

```
get_eurostat_raw(id)
```

Arguments

id A code name for the data set of interested. See the table of contents of eurostat datasets for more details.

Details

Data is downloaded from <http://ec.europa.eu/eurostat/estat-navtree-portlet-prod/BulkDownloadListing>

Value

A dataset in data.frame format. First column contains comma separated codes of cases. Other columns usually corresponds to years and column names are years with preceding X. Data is in character format as it contains values together with eurostat flags for data.

Author(s)

Przemyslaw Biecek, Leo Lahti and Janne Huovari <ropengov-forum@googlegroups.com>

References

see citation("eurostat").

See Also

[get_eurostat](#).

Examples

```
## Not run:
  tmp <- get_eurostat_raw("educ_iste")
  head(tmp)

## End(Not run)
```

get_eurostat_toc	<i>Download a table of contents of eurostat datasets.</i>
------------------	---

Description

Download table of contents (TOC) of eurostat datasets. Note that the values in column 'code' should be used to download a selected dataset.

Usage

```
get_eurostat_toc()

getEurostatTOC()
```

Details

The TOC is downloaded from http://ec.europa.eu/eurostat/estat-navtree-portlet-prod/BulkDownloadListing?sort=1&file=table_of_contents_en.txt

Value

A data.frame with eight columns

title	The name of dataset of theme
code	The codename of dataset of theme, will be used by the eurostat and get_eurostat_raw functions.
type	Is it a dataset, folder or table.
last.update.of.data, last.table.structure.change, data.start, data.end	Dates.

Functions

- getEurostatTOC: Old deprecated version

Author(s)

Przemyslaw Biecek and Leo Lahti <ropengov-forum@googlegroups.com>

References

See citation("eurostat").

See Also

[get_eurostat](#), [search_eurostat](#).

Examples

```
## Not run: tmp <- get_eurostat_toc(); head(tmp)
```

label_eurostat	<i>Get definitions for Eurostat codes from Eurostat dictionaries</i>
----------------	--

Description

A string or a factor vector of codes returns a corresponding vector of definitions. For vectors a dictionary name have to be supplied. For data.frames dictionary names are taken from column names. "time" and "values" columns are returned as they were, so you can supply data.frame from [get_eurostat](#) and get data.frame with definitions instead of codes.

Usage

```
label_eurostat(x, dic = NULL)
```

```
label_eurostat_vars(x)
```

```
label_eurostat_tables(x)
```

Arguments

x	a vector or a data.frame.
dic	a string (vector) naming eurostat dictionary or dictionaries. If NULL (default) dictionary names taken from column names of the data.frame.

Value

a vector or a data.frame.

Functions

- `label_eurostat_vars`: Get definitions for variable (column) names
- `label_eurostat_tables`: Get definitions for table names

Author(s)

Janne Huovari <janne.huovari@ptt.fi>

Examples

```
## Not run:
lp <- get_eurostat("nama_aux_lp")
lp1 <- label_eurostat_vars(lp)
str(lp1)
label_eurostat_vars(names(lp))
label_eurostat_tables("nama_aux_lp")

## End(Not run)
```

search_eurostat	<i>Lists names of dataset from eurostat with the particular pattern in the description.</i>
-----------------	---

Description

Downloads list of all datasets available on eurostat and return list of names of datasets that contains particular pattern in the dataset description. E.g. all datasets related to education of teaching.

Arguments:

Usage

```
search_eurostat(pattern, type = "dataset")
```

```
grepEurostatTOC(pattern, type = "dataset")
```

Arguments

pattern Character, datasets, folder or tables with this pattern in the description will be returned (depending on the 'type' argument)

type Grep the Eurostat table of contents either for 'dataset', 'folder' or 'table'.
Returns:

Value

A data.frame with eight columns

title The name of dataset of theme

code The codename of dataset of theme, will be used by the get_eurostat and get_eurostat_raw functions.

type Is it a dataset, folder or table.

last.update.of.data, last.table.structure.change, data.start, data.end
Dates.

Functions

- grepEurostatTOC: Old deprecated version

Author(s)

Przemyslaw Biecek and Leo Lahti <ropengov-forum@googlegroups.com>

References

See citation("eurostat")

See Also

[get_eurostat](#), [get_eurostat_toc](#)

Examples

```
## Not run: tmp <- search_eurostat("education"); head(tmp)
```

Index

*Topic **database**

- get_eurostat_dic, 6
- get_eurostat_raw, 7
- get_eurostat_toc, 8
- search_eurostat, 10

*Topic **package**

- eurostat-package, 2

*Topic **utilities**

- get_eurostat_dic, 6
- get_eurostat_raw, 7
- get_eurostat_toc, 8
- search_eurostat, 10

as.numeric, 4

clean_eurostat_cache, 3

Date, 3, 5

eurostat (eurostat-package), 2

eurostat-package, 2

eurotime2date, 3, 5

eurotime2num, 4, 5

get_eurostat, 3, 5, 7–9, 11

get_eurostat_dic, 6

get_eurostat_raw, 7

get_eurostat_toc, 8, 11

getEurostatDictionary

- (get_eurostat_dic), 6

getEurostatTOC (get_eurostat_toc), 8

grepEurostatTOC (search_eurostat), 10

label_eurostat, 6, 7, 9

label_eurostat_tables (label_eurostat),
9

label_eurostat_vars (label_eurostat), 9

search_eurostat, 7, 9, 10

tempdir, 5

tidy_eurostat, 5