# Package ‘SmarterPoland’

March 28, 2016

<table>
<thead>
<tr>
<th>Type</th>
<th>Package</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Tools for Accessing Various Datasets Developed by the Foundation SmarterPoland.pl</td>
</tr>
<tr>
<td>Version</td>
<td>1.7</td>
</tr>
<tr>
<td>Date</td>
<td>2016-03-21</td>
</tr>
<tr>
<td>Author</td>
<td>Przemyslaw Biecek (Foundation SmarterPoland.pl)</td>
</tr>
<tr>
<td>Maintainer</td>
<td>Przemyslaw Biecek <a href="mailto:przemyslaw.biecek@gmail.com">przemyslaw.biecek@gmail.com</a></td>
</tr>
<tr>
<td>Description</td>
<td>Tools for accessing and processing datasets prepared by the Foundation SmarterPoland.pl. Among all: access to API of Google Maps, Central Statistical Office of Poland, MojePanstwo, Eurostat, WHO and other sources.</td>
</tr>
<tr>
<td>LazyLoad</td>
<td>yes</td>
</tr>
<tr>
<td>LazyData</td>
<td>yes</td>
</tr>
<tr>
<td>License</td>
<td>GPL-3</td>
</tr>
<tr>
<td>Depends</td>
<td>R (&gt;= 3.0), ggplot2, httr, htmltools</td>
</tr>
<tr>
<td>Imports</td>
<td>jsonlite, rjson</td>
</tr>
<tr>
<td>NeedsCompilation</td>
<td>no</td>
</tr>
<tr>
<td>Repository</td>
<td>CRAN</td>
</tr>
<tr>
<td>Date/Publication</td>
<td>2016-03-28 13:59:24</td>
</tr>
</tbody>
</table>

R topics documented:

- SmarterPoland-package ........................................ 2
- Bank Danych Lokalnych ........................................ 3
- cities_lon_lat .................................................. 4
- countries ......................................................... 5
- Dark Sky API for weather forecast ............................ 5
- getEurostatDictionary ........................................ 7
- getEurostatRaw ................................................... 8
- getEurostatRCV .................................................. 9
- getEurostatTOC .................................................. 10
- getGoogleMapsAddress ........................................ 11
Description

Tools for accessing and processing datasets prepared by the Foundation SmarterPoland.pl. Among all: access to API of Google Maps, Central Statistical Office of Poland, Eurostat, WHO and other sources.

Details

Package: SmarterPoland
Type: Package
Version: 1.7
Date: 2016-03-21
License: GPL-3

Author(s)

Author: Przemyslaw Biecek (Foundation SmarterPoland.pl) Maintainer: Przemyslaw Biecek <przemyslaw.biecek@gmail.com>

References

http://www.smarterpoland.pl

See Also

gMillwardBrown, getEurostatRCV, getBDLseries, getWeatherForecast

Examples

```r
## Not run:
# download the dataset 'Pupil/Student - teacher ratio and average class' from eurostat
# for more developed API see https://github.com/rOpenGov/eurostat
tmp <- getEurostatRCV(kod = "educ_iste")
head(tmp)
```
Bank Danych Lokalnych

# download the dataset 'People killed in road accidents' from eurostat
# and plot a map/table for selected countries
# for more developed API see https://github.com/rOpenGov/eurostat
library(ggplot2)
t1 <- getEurostatRCV("tsdtr420")
t1 <- t1[,!\$geo]
ggplot(t1, aes(timeL valueL colorL sexL groupL sexI)) +
geom_line() + facet_wrap(~geo)

## End(Not run)

---

### Description

Access to the GUS Bank Danych Lokalnych with the use of API developed by MojePanstwo.

Download and parse data from Bank Danych Lokalnych with the use of API developed by MojePanstwo.

### Usage

```r
getBDLtree(raw = FALSE, debug = 0)
getBDLsearch(query = "", debug = 0, raw = FALSE)
getBDLseries(metric_id = "", slice = NULL, time_range = NULL,
  wojewodztwo_id = NULL, powiat_id = NULL, gmina_id = NULL,
  meta = NULL, debug = 0, raw = FALSE)
getMPgminy(debug = 0)
getMPpowiaty(debug = 0)
getMPwojewodztwa(debug = 0)
```

### Arguments

- **debug** Level of debug info. 0 for no debug, 1 or 2 for info about processed groups.
- **raw** If raw = TRUE the resulting JSON is returned without any transformation. For raw = FALSE results are transformed into a data.frame.
- **query** A query for DBL search.
- **metric_id** Metric id, if unknown then look for it in DBL tree or DBL search.
- **slice** A table with id dimensions, with format [1,34,*]. Use "*" to choose all dimensions (or use an empty string).
- **time_range** Year or range (like 2000:2010), empty means - full range.
- **wojewodztwo_id** Voievodship id or "*" for all.
- **powiat_id** County id of "*" for all. It's internal ID. Use getMPpowiaty() to get names and other information.
- **gmina_id** Subcounty id or "*" for all. It's internal ID. Use getMPgminy() to get TERYT codes.
- **meta** Should meta data be returned?
Value

The function `getMPgminy()` returns a data frame with identifiers id/TERYT for each subcounty. The function `getMPpowiaty()` returns a data frame with identifiers id for each county. The function `getBDLtree()` returns a data frame with identifiers of resources in Bank Danych Lokalnych.

Author(s)

Przemyslaw Biecek

References

The API of Bank Danych Lokalnych developed by MojePanstwo is described as https://mojepanstwo.pl/api/dane/get_

Examples

```r
## Not run:
# the data is downloaded and parsed from Internet
# not that this dataset is pre-calculated in the package
BDLtree <- getBDLtree(2)
head(BDLtree)

DBLtransport <- getBDLsearch("transport")
head(DBLtransport)

BDLseries <- getBDLseries(metric_id = 1)
head(BDLseries)

gminy <- getMPgminy()
head(gminy)

powiaty <- getMPpowiaty()
head(powiaty)

## End(Not run)
```

cities_lon_lat

---

**Geocoordinates of Largest Cities**

Description

A subset of world.citiesmaps. Extracted in order to shrink number of dependencies. Only cities with pop > 50k are kept.

Author(s)

Przemyslaw Biecek [based on world.cities]
## countries

**Birth and death rates, continent and population for selected countries**

### Description


### Author(s)

Przemyslaw Biecek [based on WHO data]

### Examples

```r
## Not run:
library(maps)
data(world.cities)
cities_lon_lat <- world.cities[!duplicated(world.cities$name),]ownames(cities_lon_lat) = cities_lon_lat[,1]
cities_lon_lat <- cities_lon_lat[cities_lon_lat$pop > 50000,]
cities_lon_lat <- cities_lon_lat[,4:5]
## End(Not run)
```

---

## Dark Sky API for weather forecast

*Access to Weather Forecasts with the Use of Dark Sky API.*

### Description

Access to hourly and daily weather forecasts with the use of Dark Sky API.

### Usage

```r
getWeatherForecast(apiKey, lat = NA, lon = NA, city = NA, raw=FALSE)
```
**Arguments**

- **apiKey**: You need to have Dark Sky apiKey in order to access weather forecasts. See here: [https://developer.forecast.io/](https://developer.forecast.io/) for more details.
- **lat**: The latitude coordinate for which prediction has to be made.
- **lon**: The longitude coordinate for which prediction has to be made.
- **city**: Instead of lat and lon you may specify name of the city for which prediction has to be made.
- **raw**: If TRUE then no parsing is done. The function `getWeatherForecast()` just download an forecast and returns it as a list.

**Value**

The function `getWeatherForecast()` returns list of three datasets. `now` and `by.hour` datasets contains predictions. For each timepoint following information are collected:

- time, summary, icon, precipIntensity, precipProbability, temperature, apparentTemperature, dewPoint, humidity, windSpeed, windBearing, visibility, cloudCover, pressure, ozone, temperatureCelsius, apparentTemperatureCelsius

Daily predictions (`by.day` component) contain following information:

- time, summary, icon, sunriseTime, sunsetTime, moonPhase, precipIntensity, precipIntensityMax, precipProbability, temperatureMin, temperatureMinTime, temperatureMax, temperatureMaxTime, apparentTemperatureMin, apparentTemperatureMinTime, apparentTemperatureMax, apparentTemperatureMaxTime, dewPoint, humidity, windSpeed, windBearing, visibility, cloudCover, pressure, ozone, precipIntensityMaxTime, precipType, temperatureMaxCelsius, temperatureMinCelsius, apparentTemperatureMaxCelsius, apparentTemperatureMinCelsius

**Author(s)**

Przemyslaw Biecek

**References**

The Dark Sky API for weather forecasts is described as [https://developer.forecast.io/](https://developer.forecast.io/)

**Examples**

```r
## Not run:
# you have to have apiKey to execute these examples
library(scales)
library(ggplot2)

prognoza <- getWeatherForecast(apiKey, city='Warsaw')

ggplot(prognoza$by.hour, aes(y=temperatureCelsius, x=time)) +
  geom_line() + geom_point() +
  geom_point(data=prognoza$now, size=10, color='red') +
  theme(title=element_text(size=20),
        axis.text=element_text(size=20)) +
  scale_x_datetime(breaks = date_breaks("3 hour"),
```


getEurostatDictionary

    minor_breaks = date_breaks("1 hour"),
    labels = date_format("y") + xlab("") + ggtitle("Prognoza temperatury dla Warszawy")

    ## End(Not run)

gEurostatDictionary  Download a Dictionary from the Eurostat Database

Description

Download a dictionary for given coded variable from Eurostat (ec.europa.eu/eurostat).

Usage

gEurostatDictionary(dictname)

Arguments

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

Value

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

Author(s)

Przemyslaw Biecek

References

The TOC is downloaded from the http://ec.europa.eu/eurostat/estat-navtree-portlet-prod/BulkDownloadListing

See Also

See Also as getEurostatRCV, getEurostatRaw, grepEurostatTOC.

Examples

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

    ## End(Not run)
getEurostatRaw

Download a Dataset from the Eurostat Database

Description
Download a dataset from the eurostat database. The dataset is transformed into the tabular format.

Usage
getEurostatRaw(kod = "educ_iste", rowRegExp=NULL, colRegExp=NULL, strip.white = TRUE)

Arguments
- **kod**: A code name for the data set of interested. See the table of contents of eurostat datasets for more details.
- **rowRegExp**: If not NULL this regular expression will be used to filter rows out of downloaded file.
- **colRegExp**: If not NULL this regular expression will be used to filter columns out of downloaded file.
- **strip.white**: Passed to the internal `read.table()`. By default it strips white spaces from eurostat values.

Value
A dataset in data.frame format. First column contains names of cases. Column names usually corresponds to years.

Author(s)
Przemyslaw Biecek

References

See Also
See Also as `getEurostatTOC, getEurostatRaw, grepEurostatTOC`.

Examples
```r
## Not run:
tmp <- getEurostatRaw(kod = "educ_iste")
head(tmp)
## End(Not run)
```
getEurostatRCV  Download a Dataset from the Eurostat Database

Description
Download a dataset from the eurostat database. The dataset is transformed into the molten / row-column-value format (RCV).

Usage
getEurostatRCV(kod = "educ_iste", ...)

Arguments
kod A code name for the data set of interested. See the table of contents of eurostat datasets for more details.
... Other parameters that are passed to getEurostatRaw().

Value
A dataset in the molten format with the last column ‘value’.

Author(s)
Przemyslaw Biecek

References

See Also
See Also as getEurostatTOC, getEurostatRaw, grepEurostatTOC.

Examples
```R
## Not run:
tmp <- getEurostatRCV(kod = "educ_iste")
head(tmp)

t1 <- getEurostatRCV("rail_ac_catvict")
tmp <- cast(t1, geo ~ time, mean, subset=victim="KIL" & pers_inv="TOTAL" & accident="TOTAL")
tmp3 <- tmp[,1:9]
rownames(tmp3) <- tmp3[,1]
tmp3 <- tmp3[c("UK", "SK", "FR", "PL", "ES", "PT", "LV"),]
matplot(2005:2012,t(tmp3,-1)), type="o", pch=19, lty=1,
las=1, xlab="", ylab="", yaxt="n")
```
getEurostatTOC

Description

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

Usage

gEurostatTOC()

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 getEurostatRCV and getEurostatRaw functions.
type Is it a dataset, folder or table.
last.update.of.data, last.table.structure.change, data.start, data.end Dates.

Author(s)

Przemyslaw Biecek

References

The TOC is downloaded from the http://ec.europa.eu/eurostat/estat-navtree-portlet-prod/BulkDownloadListing

See Also

See Also as getEurostatRCV, getEurostatRaw, grepEurostatTOC.

Examples

## Not run:
tmp <- getEurostatTOC()
head(tmp)

## End(Not run)
Description

Get geolocation (longitude, latitude) of a given address with the use of Google Maps API.

The Google Maps API is used to determine the geolocation (longitude, latitude) of a given address.

Usage

```r
googleMapsAddress(street = "Banacha 2", city = "Warszawa", country="Poland", positionOnly = TRUE, delay=1)
```

Arguments

- **street**: An address (street and building number)
- **city**: City
- **country**: Country
- **positionOnly**: What should be returned, vector with longitude, latitude coordinates or the raw result from Google Maps API
- **delay**: Number of seconds to wait between api calls

Value

If positionOnly=TRUE then a vector with two values or a raw list from Google Maps otherwise.

Author(s)

Przemyslaw Biecek

References

The Google Maps API https://developers.google.com/maps/

Examples

```r
## Not run:
googleMapsAddress()

## End(Not run)```
getMillwardBrown  

*MillwardBrown Pool Results*

**Description**

Download pool results from MillwardBrown website.

**Usage**

```r
getMillwardBrown()
```

**Value**

A dataset in the molten format with pool date, party and percent of votes.

**Author(s)**

Maciej Beresewicz [data extraction] Przemyslaw Biecek [data melting]

**Examples**

```r
# Not run:
getMillwardBrown()

# End(Not run)
```

---

grepEurostatTOC  

*Names of Eurostat Datasets That Fit Given Pattern*

**Description**

Lists names of dataset from eurostat with the particular pattern in the description.

This function 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.

**Usage**

```r
grepEurostatTOC(pattern)
```

**Arguments**

- `pattern`  
  Character, only datasets that contains this pattern in the description will be returned.
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 getEurostatRCV and getEurostatRaw functions.
type                  Is it a dataset, folder or table.
last.update.of.data, last.table.structure.change, data.start, data.end Dates.

Author(s)

Przemyslaw Biecek

See Also

See Also as getEurostatRCV, getEurostatRaw, getEurostatTOC.

Examples

## Not run:
tmp <- grepEurostatTOC("education")
head(tmp)

## End(Not run)

---

maturaExam   Results from Matura Exams in Poland for Math and Language for Years 2010-2015

Description

This dataset is created based on data from ZPD package, see https://github.com/zozlak/ZPD and http://zpd.ibe.edu.pl/doku.php?id=obazie. Each row shows results for one person that takes matura exams in a given year.

Author(s)

Przemyslaw Biecek [based on IBE / ZPD data]

Examples

## Not run:
data(maturaExam)
head(maturaExam)

## End(Not run)
Index

*Topic database
  Bank Danych Lokalnych, 3
  Dark Sky API for weather forecast, 5
    getEurostatDictionary, 7
    getEurostatRaw, 8
    getEurostatRCV, 9
    getEurostatTOC, 10
    grepEurostatTOC, 12

*Topic datasets
  cities_lon_lat, 4
  countries, 5
  maturaExam, 13

*Topic geolocalisation
  getGoogleMapsAddress, 11

*Topic package
  SmarterPoland-package, 2

Bank Danych Lokalnych, 3
BDLtree (Bank Danych Lokalnych), 3

cities_lon_lat, 4
countries, 5

Dark Sky API for weather forecast, 5

getBDLsearch (Bank Danych Lokalnych), 3
getBDLseries, 2
getBDLseries (Bank Danych Lokalnych), 3
getBDLtree (Bank Danych Lokalnych), 3
getEurostatDictionary, 7
getEurostatRaw, 7, 8, 9, 10, 13
getEurostatRCV, 2, 7, 9, 10, 13
getEurostatTOC, 8, 9, 10, 13
getGoogleMapsAddress, 11
getStatus, 11
getMillwardBrown, 2, 12
getMPgminy (Bank Danych Lokalnych), 3
getMPpowiats (Bank Danych Lokalnych), 3
getMPwojvodzta (Bank Danych Lokalnych), 3