

Package ‘insee’

August 3, 2020

Type Package

Title Tools to Easily Download Data from INSEE BDM Database

Version 0.8.0

Author Hadrien Leclerc [aut, cre]

Maintainer Hadrien Leclerc <leclerc.hadrien@gmail.com>

Description Using embedded sdmx queries, get the data of more than 140 000 insee series from bdm database. Have a look at the detailed sdmx web service page with the following link : <<https://www.insee.fr/en/information/2868055>>.

URL <<https://www.insee.fr/en/information/2868055>>

Encoding UTF-8

LazyData true

License GPL-2

Imports rsdmx, httr, xml2, tibble, dplyr, tidyr, stringr, lubridate, tidyselect, rlang

Suggests tidyverse, prettydoc, htmltools, kableExtra

RoxygenNote 7.1.1

NeedsCompilation no

Repository CRAN

Date/Publication 2020-08-02 22:10:02 UTC

R topics documented:

get_dataset_list	2
get_idbank_list	2
get_insee	3
get_insee_dataset	3
get_insee_idbank	4
get_insee_title	6

Index	8
--------------	----------

get_dataset_list	<i>Download an INSEE dataset list</i>
------------------	---------------------------------------

Description

Download an INSEE dataset list

Usage

```
get_dataset_list()
```

Details

the datasets returned are the ones available through a SDMX query

Examples

```
insee_dataset = get_dataset_list()
```

get_idbank_list	<i>Download a mapping dataset between INSEE series keys (idbank) and SDMX series names</i>
-----------------	--

Description

Download a mapping dataset between INSEE series keys (idbank) and SDMX series names

Usage

```
get_idbank_list(dataset = NULL)
```

Arguments

dataset	if a dataset name is provided, only a subset of the data is delivered, otherwise all the data is returned
---------	---

Details

Download a mapping dataset between INSEE series keys (idbank) and SDMX series names

Examples

```
idbank_list = get_idbank_list()
```

`get_insee`*Get data from INSEE BDM database with a SDMX query link*

Description

Get data from INSEE BDM database with a SDMX query link

Usage

```
get_insee(link)
```

Arguments

`link` SDMX query link

Details

Get data from INSEE BDM database with a SDMX query link Mainly for package internal use from the functions `get_insee_dataset` and `get_insee_idbank`

Examples

```
insee_link = "http://www.bdm.insee.fr/series/sdmx/data/SERIES_BDM"  
insee_query = file.path(insee_link, paste0("010539365","?", "firstNObservations=1"))  
data = get_insee(insee_query)
```

`get_insee_dataset`*Get dataset from INSEE BDM database*

Description

Get dataset from INSEE BDM database

Usage

```
get_insee_dataset(  
  dataset,  
  startPeriod = NULL,  
  endPeriod = NULL,  
  firstNObservations = NULL,  
  lastNObservations = NULL,  
  filter = NULL  
)
```

Arguments

dataset	dataset name to be downloaded
startPeriod	start date of data
endPeriod	end date of data
firstNObservations	get the first N observations for each key series (idbank)
lastNObservations	get the last N observations for each key series (idbank)
filter	Use the filter to choose only some dimensions. It is recommended to use it for big datasets. A dimension left empty means all values are selected. To select multiple values in one dimension put a "+" between those values (see example)

Details

Get dataset from INSEE BDM database

Examples

```
insee_dataset = get_dataset_list()

data = get_insee_dataset("IPC-2015", filter = "M+A.....CVS.", startPeriod = "2015-03")
```

get_insee_idbank	<i>Get data from INSEE series idbank</i>
------------------	--

Description

Get data from INSEE series idbank

Usage

```
get_insee_idbank(
  ...,
  startPeriod = NULL,
  endPeriod = NULL,
  firstNObservations = NULL,
  lastNObservations = NULL
)
```

Arguments

... one or several series key (idbank)
 startPeriod start date of data
 endPeriod end date of data
 firstNObservations
 get the first N observations for each key series (idbank)
 lastNObservations
 get the last N observations for each key series (idbank)

Details

Get data from INSEE series idbank

Examples

```
#example 1 : import price index of industrial products and turnover index : manufacture of wood
data = get_insee_idbank("001558315", "010540726")
```

```
#example 2 : unemployment data
```

```
library(tidyverse)
```

```
idbank_list = get_idbank_list()
```

```
df_idbank_list_selected =
  idbank_list %>%
  filter(nomflow == "CHOMAGE-TRIM-NATIONAL") %>% #unemployment dataset
  filter(dim5 == 0) %>% #men and women
  mutate(title = get_insee_title(idbank))
```

```
idbank_list_selected = df_idbank_list_selected %>% pull(idbank)
```

```
unem = get_insee_idbank(idbank_list_selected)
```

```
#example 3 : French GDP growth rate
```

```
library(tidyverse)
```

```
idbank_list = get_idbank_list()
```

```
df_idbank_list_selected =
  idbank_list %>%
  filter(nomflow == "CNT-2014-PIB-EQB-RF") %>% # Gross domestic product balance
  filter(dim1 == "T") %>% #quarter
  filter(dim4 == "PIB") %>% #GDP
  filter(dim6 == "TAUX") %>% #rate
  filter(dim10 == "CVS-CJ0") #SA-WDA, seasonally adjusted, working day adjusted
```

```
idbank = idbank_list_selected %>% pull(idbank)
```

```

data = get_insee_idbank(idbank)

#plot
ggplot(data, aes(x = DATE, y = OBS_VALUE)) +
  geom_col() +
  ggtitle("French GDP growth rate, quarter-on-quarter, sa-wda") +
  labs(subtitle = sprintf("Last updated : %s", data$TIME_PERIOD[1]))

```

get_insee_title	<i>Get title from INSEE series idbank</i>
-----------------	---

Description

Get title from INSEE series idbank

Usage

```
get_insee_title(..., lang = "en")
```

Arguments

...	list of series key (idbank)
lang	language of the title, by default it is English, if lang is different from "en" then French will be the title's language

Details

Query INSEE website to get series title from series key (idbank). Any query to INSEE database can handle around 400 idbanks at maximum, if necessary the idbank list will then be splitted in several lists of 400 idbanks each. Consequently, it is not advised to use it on the whole idbank dataset, the user should filter the idbank dataset first.

Examples

```
#example 1 : industrial production index on manufacturing and industrial activities
title = get_insee_title("010537900")
```

```
#example 2 : automotive industry and overall industrial production
library(tidyverse)
```

```
idbank_list = get_idbank_list()
```

```
idbank_list_selected =
  idbank_list %>%
  filter(nomflow == "IPI-2015") %>% #industrial production index dataset
```

```
filter(dim1 == "M") %>% #monthly
filter(dim5 == "INDICE") %>% #index
filter(dim8 == "CVS-CJ0") %>% #Working day and seasonally adjusted SA-WDA
filter(str_detect(dim4, "^29$|A10-BE")) %>% #automotive industry and overall industrial production
mutate(title = get_insee_title(idbank))
```

Index

`get_dataset_list`, 2
`get_idbank_list`, 2
`get_insee`, 3
`get_insee_dataset`, 3
`get_insee_idbank`, 4
`get_insee_title`, 6