

Package ‘datazoom.amazonia’

October 13, 2022

Title Simplify Access to Data from the Amazon Region

Version 0.3.0

Description Functions to download and treat data regarding the Brazilian Amazon region from a variety of official sources.

License MIT + file LICENSE

URL <http://www.econ.puc-rio.br/datazoom/>

Depends R (>= 4.0)

Imports data.table, dplyr, Hmisc, janitor, labelled, lubridate, magrittr, purrr, RCurl, readr, readxl, sf, sidrar, stats, stringi, stringr, tibble, tidyr, utils, XML

Suggests googledrive, knitr, rmarkdown, terra

VignetteBuilder knitr

Encoding UTF-8

RoxygenNote 7.1.2

NeedsCompilation no

Author Francisco de Lima Cavalcanti [aut, cre],
DataZoom (PUC-Rio) [fnd],
Guilherme Jardim [aut],
Daniel AC Barbosa [aut],
Bruno Alcantara Duarte [aut],
Fredie Didier [aut],
Tito Bruni [aut],
Victor Aliende da Matta [aut],
Anna Carolina Dutra Saraiva [aut],
Arthur Carvalho Brito [aut],
Igor Rigolon Veiga [aut],
Maria Mittelbach [aut]

Maintainer Francisco de Lima Cavalcanti <francisco.lima.cavalcanti@gmail.com>

Repository CRAN

Date/Publication 2022-04-08 08:12:29 UTC

R topics documented:

| | |
|--------------------------|-----------|
| load_baci | 2 |
| load_br_trade | 3 |
| load_cempre | 4 |
| load_cipo | 5 |
| load_climate | 6 |
| load_degrad | 7 |
| load_deter | 8 |
| load_ibama | 9 |
| load_ips | 10 |
| load_mapbiomas | 11 |
| load_pam | 12 |
| load_pevs | 13 |
| load_pibmunic | 14 |
| load_ppm | 15 |
| load_prodes | 16 |
| load_seeg | 17 |
| load_sigmine | 18 |
| Index | 19 |

| | |
|-----------|------------------------------|
| load_baci | <i>BACI - External Trade</i> |
|-----------|------------------------------|

Description

provides disaggregated data on bilateral trade flows for more than 5000 products and 200 countries.
See http://www.cepii.fr/CEPII/en/bdd_modele/presentation.asp?id=37.

Usage

```
load_baci(dataset = "HS92", raw_data, time_period, language = "pt")
```

Arguments

| | |
|-------------|--|
| dataset | A dataset name ("HS92"). |
| raw_data | A boolean setting the return of raw (TRUE) or processed (FALSE) data. |
| time_period | A numeric indicating what years will the data be loaded in the format YYYY. Can be only one year at a time. |
| language | A string that indicates in which language the data will be returned. Currently, only Portuguese ("pt") and English ("eng") are supported. Defaults to "pt". |

Value

A tibble consisting of imports or exports data.

Examples

```
## Not run:
# download treated trade data from 2018
exp_mun <- load_br_trade(
  dataset = "HS92",
  raw_data = FALSE, time_period = 2018
)

## End(Not run)
```

load_br_trade

Comex - External Trade

Description

Loads information on both imports and exports data. Data is available from 1997 to 2021 for most datasets. See <https://www.gov.br/produtividade-e-comercio-exterior/pt-br/assuntos/comercio-exterior/estatisticas/base-de-dados-bruta/>.

Usage

```
load_br_trade(dataset = NULL, raw_data, time_period, language = "eng")
```

Arguments

| | |
|-------------|--|
| dataset | A dataset name ("comex_export_mun", "comex_import_mun", "comex_export_prod" or "comex_import_prod"). |
| raw_data | A boolean setting the return of raw (TRUE) or processed (FALSE) data. |
| time_period | A numeric indicating what years will the data be loaded in the format YYYY. Can be a sequence of numbers such as 2010:2012. |
| language | A string that indicates in which language the data will be returned. Currently, only Portuguese ("pt") and English ("eng") are supported. Defaults to "eng". |

Value

A tibble consisting of imports or exports data.

Examples

```
## Not run:
# download treated exports data by municipality from 1997 to 2021
exp_mun <- load_br_trade(
  dataset = "comex_export_mun",
  raw_data = FALSE, time_period = 1997:2021
)

# download raw imports data by municipality from 1997 to 2021
```

```

raw_imp_mun <- load_br_trade(
  dataset = "comex_import_mun",
  raw_data = TRUE, time_period = 1997:2021
)

# download treated imports data by ncm from 1997 to 2021
imp_prod <- load_br_trade(
  dataset = "comex_import_prod",
  raw_data = FALSE, time_period = 1997:2021
)

## End(Not run)

```

load_cempre

CEMPRE - Central Register of Companies

Description

Loads information on companies and other organizations and their respective formally constituted local units, registered with the CNPJ - National Register of Legal Entities. Data is available from 2006 to 2019. See <https://sidra.ibge.gov.br/pesquisa/empre/tabelas>

Usage

```

load_cempre(
  dataset = "cempre",
  raw_data,
  geo_level,
  time_period,
  language = "eng",
  sectors = FALSE,
  legal_amazon_only = FALSE
)

```

Arguments

| | |
|-------------|--|
| dataset | A dataset name ("cempre"). |
| raw_data | A boolean setting the return of raw (TRUE) or processed (FALSE) data. |
| geo_level | A string that defines the geographic level of the data. Can be one of "country", "state" or "municipality". See documentation of <code>sidrar</code> . |
| time_period | A numeric indicating what years will the data be loaded in the format YYYY. Can be a sequence of numbers such as 2010:2012. |
| language | A string that indicates in which language the data will be returned. Currently, only Portuguese ("pt") and English ("eng") are supported. Defaults to "eng". |
| sectors | A boolean that defines if the data will be return separated by sectors (TRUE) or not (FALSE). Defaults to FALSE |

```
legal_amazon_only
```

A boolean setting the return of Legal Amazon Data (TRUE) or Country's Data (FALSE). Defaults to FALSE

Value

A tibble with the selected data.

Examples

```
## Not run:  
# download raw data from 2006 to 2019  
raw_cempre_all <- load_cempre(  
  dataset = "cempre",  
  raw_data = TRUE,  
  geo_level = "municipality",  
  time_period = 2006:2019  
)  
  
## End(Not run)
```

load_cipo

Plataforma CIPÓ - Mappings on environmental crimes

Description

Plataforma CIPÓ - Mappings on environmental crimes

Usage

```
load_cipo(dataset = "brazilian_actors", search = "")
```

Arguments

dataset A dataset name ("brazilian_actors", "international_cooperation", "forest_governance")
search A string that filters entries containing it.

Value

A tibble of the chosen CIPÓ spreadsheet.

Examples

```
## Not run:
# download the spreadsheet on Brazilian actors involved in fighting environmental crimes
brazilian_actors <- load_cipo(dataset = "brazilian_actors")

# searching only for entries containing IBAMA
actors_ibama <- load_cipo(
  dataset = "brazilian_actors",
  search = "ibama"
)

# entries containing IBAMA or FUNAI
actors_ibama <- load_cipo(
  dataset = "brazilian_actors",
  search = "ibama|funai"
)

## End(Not run)
```

load_climate

TerraClimate

Description

TerraClimate

Usage

```
load_climate(
  dataset,
  raw_data = FALSE,
  time_period,
  language = "eng",
  legal_amazon_only = FALSE
)
```

Arguments

| | |
|-------------|---|
| dataset | A dataset name, choosing which variable will be loaded. One of ("max_temperature", "min_temperature", "wind_speed", "vapor_pressure_deficit", "vapor_pressure", "snow_water_equivalent", "shortwave_radiation_flux", "soil_moisture", "runoff", "precipitation", "potential_evaporation", "climatic_water_deficit", "water_evaporation", "palmer_drought_severity_index"). For extra details, try vignette("TERRACLIMATE"). |
| raw_data | A boolean setting the return of raw (TRUE) or processed (FALSE) data. |
| time_period | A numeric indicating for what years the data will be loaded in the format YYYY. Can be a sequence of numbers such as 2010:2012. |
| language | A string that indicates in which language the data will be returned. Currently, only Portuguese ("pt") and English ("eng") are supported. Defaults to "eng". |

legal_amazon_only

A boolean setting the return of Legal Amazon Data (TRUE) or Country's Data (FALSE). Defaults to FALSE

Value

A tibble

Examples

```
## Not run:  
max_temp <- load_climate(dataset = "max_temperature", time_period = 2000:2020)  
  
## End(Not run)
```

load_degrad

Degrad - Forest Degradation in the Brazilian Amazon

Description

Loads information on forest degradation in the Brazilian Amazon, replaced by DETER-B in December 2016. Data is available from 2007 to 2016. See <http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/degrad>.

Usage

```
load_degrad(dataset = "degrad", raw_data, time_period, language = "eng")
```

Arguments

| | |
|-------------|--|
| dataset | A dataset name ("degrad"). |
| raw_data | A boolean setting the return of raw (TRUE) or processed (FALSE) data. |
| time_period | A numeric indicating what years will the data be loaded in the format YYYY. Can be a sequence of numbers such as 2010:2012. |
| language | A string that indicates in which language the data will be returned. Currently, only Portuguese ("pt") and English ("eng") are supported. Defaults to "eng". |

Value

A list (if raw_data = TRUE) or a tibble (if raw_data = FALSE) with the selected data.

Examples

```
## Not run:
# download raw data (raw_data = TRUE) related to forest degradation
# from 2010 to 2012 (time_period = 2010:2012).
data <- load_degrad(
  dataset = "degrad",
  raw_data = TRUE,
  time_period = 2010:2012
)

# download treated data (raw_data = FALSE) related to forest degradation
# from 2013 (time_period = 2013) in portuguese (language = "pt").
data <- load_degrad(
  dataset = "degrad",
  raw_data = FALSE,
  time_period = 2013,
  language = "pt"
)

## End(Not run)
```

load_deter

DETER - Forest Degradation in the Brazilian Amazon

Description

Loads information on change in forest cover in the Amazon. See <http://www.obt.inpe.br/OBT/assuntos/programas/amazonia/deter/deter>

Usage

```
load_deter(dataset = NULL, raw_data, language = "eng")
```

Arguments

| | |
|----------|--|
| dataset | A dataset name ("deter_amz", "deter_cerrado") with information about both Amazon and Cerrado |
| raw_data | A boolean setting the return of raw (TRUE) or processed (FALSE) data. |
| language | A string that indicates in which language the data will be returned. Currently, only Portuguese ("pt") and English ("eng") are supported. Defaults to "eng". |

Value

A tibble with the selected data.

Examples

```
## Not run:
# Download raw data (raw_data = TRUE) from Amazonia (dataset = "deter_amz")
deter_amz <- load_deter(
  dataset = "deter_amz",
  raw_data = TRUE
)

# Download treated data (raw_data = FALSE) from Cerrado (dataset = "deter_cerrado")
# in portuguese (language = 'pt')
deter_cer <- load_deter(
  dataset = "deter_cerrado",
  raw_data = FALSE,
  language = "pt"
)

## End(Not run)
```

load_ibama

IBAMA - Brazilian Institute for the Environment and Renewable Natural Resources

Description

Loads information on on environmental fines at the municipality or state levels considering the Amazon region

Usage

```
load_ibama(
  dataset = "areas_embargadas",
  raw_data,
  geo_level = "municipality",
  language = "eng",
  legal_amazon_only = FALSE
)
```

Arguments

| | |
|-------------------|---|
| dataset | A dataset name ("areas_embargadas") |
| raw_data | A boolean setting the return of raw or processed data |
| geo_level | A string that defines the geographic level of the data. Can be only "municipality". |
| language | A string that indicates in which language the data will be returned. Currently, only Portuguese ("pt") and English ("eng") are supported. |
| legal_amazon_only | A boolean setting the return of Legal Amazon Data or not |

Value

A tibble with the selected data.

Examples

```
## Not run:  
# download raw data from all country  
raw_ibama_all <- load_ibama(  
  dataset = "areas_embargadas",  
  raw_data = TRUE,  
  legal_amazon_only = FALSE  
)  
  
## End(Not run)
```

load_ips

IPS - Amazon Social Progress Index

Description

Loads information on the social and environmental performance of the Legal Amazon. Survey is done at the municipal level and data is available in 2014 and 2018. See <http://www.ipsamazonia.org.br/>

Usage

```
load_ips(dataset = "ips", raw_data, time_period, language = "eng")
```

Arguments

| | |
|-------------|--|
| dataset | A dataset name ("ips") |
| raw_data | A boolean setting the return of raw or processed data |
| time_period | A numeric indicating what years will the data be loaded in the format YYYY. |
| language | A string that indicates in which language the data will be returned. Currently, only Portuguese and English are supported. |

Value

A tibble with the selected data.

Examples

```
## Not run:  
# download raw data from 2014  
ips <- load_ips(dataset = "ips", raw_data = TRUE, time_period = 2014)  
  
## End(Not run)
```

| | |
|----------------|--|
| load_mapbiomas | <i>MAPBIOMAS - The Annual Land Cover and Use Mapping Project in Brazil</i> |
|----------------|--|

Description

Loads information about land cover and use

Usage

```
load_mapbiomas(
  dataset = NULL,
  raw_data = NULL,
  geo_level = "municipality",
  time_period = "all",
  language = "eng",
  time_id = "year",
  cover_level = 1
)
```

Arguments

| | |
|-------------|---|
| dataset | A dataset name ("mapbiomas_cover", "mapbiomas_transition", "mapbiomas_irrigation", "mapbiomas_deforestation_regeneration", "mapbiomas_grazing_quality") |
| raw_data | A boolean setting the return of raw or processed data |
| geo_level | A string that defines the geographic level of the data. Can be only "municipality". |
| time_period | A numeric indicating what years will the data be loaded. Can be only "all". |
| language | A string that indicates in which language the data will be returned. Currently, only Portuguese ("pt") and English ("eng") are supported. |
| time_id | A string that indicates the time criteria for the data loaded. Can be "year" or "month". Defaults to year. |
| cover_level | A numeric that indicates the cover aggregation level. Can be "0", "1", "2", "3" or "4". |

Value

A tibble with the selected data.

Examples

```
## Not run:
# download treated data from mapbiomas_grazing_quality
treated_mapbiomas_grazing <- load_mapbiomas(
  dataset = "mapbiomas_grazing_quality",
  raw_data = FALSE, geo_level = "municipality",
```

```

  time_period = "all", language = "pt"
)

## End(Not run)

```

load_pam

PAM - Municipal Agricultural Production

Description

Loads information on the temporary and permanent crops of the Country that are characterized not only by their great economic importance in the export agenda, but also by their social relevance, since its components are on the Brazilian table (IBGE). Survey is done at the municipal level and data is available from 1974 to 2020 for most datasets. See <https://www.ibge.gov.br/en/statistics/economic/agriculture-forestry-and-fishing/16773-municipal-agricultural-production-tempor.html?=&t=o-que-e>

Usage

```
load_pam(dataset = NULL, raw_data, geo_level, time_period, language = "eng")
```

Arguments

| | |
|-------------|--|
| dataset | A dataset name ("all_crops", "permanent_crops", "temporary_crops" or many individual crop possibilities (see <code>vignette(load_pam)</code>)). You can also use SIDRA codes (see https://sidra.ibge.gov.br/pesquisa/pam/tabelas) |
| raw_data | A boolean setting the return of raw (TRUE) or processed (FALSE) data. |
| geo_level | A string that defines the geographic level of the data. Can be one of "country", "state" or "municipality". See documentation of <code>sidrar</code> . |
| time_period | A numeric indicating what years will the data be loaded in the format YYYY. Can be a sequence of numbers such as 2010:2012. |
| language | A string that indicates in which language the data will be returned. Currently, only Portuguese ("pt") and English ("eng") are supported. Defaults to "eng". |

Value

A tibble consisting of geographic units that present positive values for any of the variables in the dataset.

Examples

```

## Not run:
# download state raw data from 2012 for all crops
pam_all_crops <- load_pam(
  dataset = "all_crops",
  raw_data = TRUE,

```

```

    geo_level = "state",
    time_period = 2012
  )

  ## End(Not run)

```

load_pevs

PEVS - Forestry Activities

Description

Loads information on the amount and value of the production of the exploitation of native plant resources and planted forest massifs, as well as existing total and harvested areas of forest crops. Survey is done at the municipal level and data is available from 1986 to 2020. See <https://www.ibge.gov.br/en/statistics/economic/agriculture-forestry-and-fishing/18374-forestry-activities.html?=&t=o-que-e>

Usage

```
load_pevs(dataset = NULL, raw_data, geo_level, time_period, language = "eng")
```

Arguments

| | |
|-------------|--|
| dataset | A dataset name ("pevs_forest_crops", "pevs_silviculture" or "pevs_silviculture_area"). You can also use SIDRA codes (see https://sidra.ibge.gov.br/pesquisa/pevs/quadros/brasil/2019) |
| raw_data | A boolean setting the return of raw (TRUE) or processed (FALSE) data. |
| geo_level | A string that defines the geographic level of the data. Can be one of "country", "region", "state", "mesoregion", "microregion" or "city". See documentation of <code>sidrar</code> . |
| time_period | A numeric indicating what years will the data be loaded in the format YYYY. Can be a sequence of numbers such as 2010:2012. |
| language | A string that indicates in which language the data will be returned. Currently, only Portuguese ("pt") and English ("eng") are supported. Defaults to "eng". |

Value

A tibble consisting of geographic units that present positive values for any of the variables in the dataset.

Examples

```

## Not run:
# download state raw data from 2012 for silviculture
pevs_silvi <- load_pevs(
  dataset = "pevs_silviculture",
  raw_data = TRUE,

```

```

    geo_level = "state",
    time_period = 2012
  )

  ## End(Not run)

```

load_pibmunic

PIB MUNICIPAL - Municipal GDP

Description

Loads information on gross domestic product at current prices, taxes, net of subsidies, on products at current prices and gross value added at current prices, total and by economic activity, and respective shares. Survey is done at Country, state and municipality level and data is available from 2002 to 2019.

Usage

```

load_pibmunic(
  dataset = "pibmunic",
  raw_data,
  geo_level,
  time_period,
  language = "eng",
  legal_amazon_only = FALSE
)

```

Arguments

| | |
|-------------------|---|
| dataset | A dataset name ("pibmunic") with Municipal GDP information. You can also use SIDRA codes (See https://sidra.ibge.gov.br/pesquisa/pib-munic/tabelas) |
| raw_data | A boolean setting the return of raw (TRUE) or processed (FALSE) data. |
| geo_level | A string that defines the geographic level of the data. Can be one of "country", "state" or "municipality". |
| time_period | A numeric indicating what years will the data be loaded in the format YYYY. Can be a sequence of numbers such as 2010:2012. |
| language | A string that indicates in which language the data will be returned. Currently, only Portuguese ("pt") and English ("eng") are supported. Defaults to "eng". |
| legal_amazon_only | A boolean setting the return of Legal Amazon Data (TRUE) or Country's Data (FALSE). |

Value

A tibble with the selected data.

Examples

```
## Not run:
# download state raw data from 2012 for all country
pibmunic <- load_pibmunic(
  dataset = "pibmunic",
  raw_data = TRUE,
  geo_level = "state",
  time_period = 2012,
  legal_amazon_only = FALSE
)

## End(Not run)
```

load_ppm

*PPM - Municipal Livestock Production***Description**

Loads information on animal farming inventories and livestock products (IBGE). Survey is done at the municipal level and data is available from 1974 to 2020 for most datasets. See <https://www.ibge.gov.br/en/statistics/economic/agriculture-forestry-and-fishing/17353-municipal-livestock-production.html?=&t=o-que-e>

Usage

```
load_ppm(dataset = NULL, raw_data, geo_level, time_period, language = "eng")
```

Arguments

| | |
|-------------|---|
| dataset | A dataset name ("ppm_livestock_inventory", "ppm_sheep_farming", "ppm_animal_orig_production", "ppm_cow_farming" or "ppm_aquaculture". You can also use SIDRA codes (see https://sidra.ibge.gov.br/pesquisa/ppm/tabelas/brasil/2019) |
| raw_data | A boolean setting the return of raw (TRUE) or processed (FALSE) data. |
| geo_level | A string that defines the geographic level of the data. Can be one of "country", "state" or "municipality". See documentation of <code>sidrar</code> . |
| time_period | A numeric indicating what years will the data be loaded in the format YYYY. Can be a sequence of numbers such as 2010:2012. |
| language | A string that indicates in which language the data will be returned. Currently, only Portuguese ("pt") and English ("eng") are supported. Defaults to "eng". |

Value

A tibble consisting of geographic units that present positive values for any of the variables in the dataset.

Examples

```
## Not run:
# download state raw data from 2012 for animal origin production
ppm_aop <- load_ppm(
  dataset = "ppm_animal_origin_production",
  raw_data = TRUE,
  geo_level = "state",
  time_period = 2012
)

## End(Not run)
```

| | |
|-------------|---|
| load_prodes | <i>PRODES - Deforestation Monitoring Project in the Legal Amazon by Satellite</i> |
|-------------|---|

Description

Loads information on clearcut deforestation in the Legal Amazon and annual deforestation rates in the region. Survey is done at state or municipality level and data is available from 2000 to 2020.

Usage

```
load_prodes(dataset = "prodes", raw_data, time_period, language = "eng")
```

Arguments

| | |
|-------------|--|
| dataset | A dataset name ("prodes"). |
| raw_data | A boolean setting the return of raw (TRUE) or processed (FALSE) data. |
| time_period | A numeric indicating what years will the data be loaded in the format YYYY. Can be a sequence of numbers such as 2010:2012. |
| language | A string that indicates in which language the data will be returned. Currently, only Portuguese ("pt") and English ("eng") are supported. Defaults to "eng". |

Value

A tibble with the selected data.

Examples

```
## Not run:
# download raw data from 2000 to 2020
raw_prodes_all <- load_prodes(
  dataset = "prodes",
  raw_data = TRUE,
  time_period = 2000:2020
)
```



```
## End(Not run)
```

| | |
|-----------|---|
| load_seeg | <i>Greenhouse gases emission estimates (SEEG)</i> |
|-----------|---|

Description

Loads data of estimates of emission of greenhouse gases

Usage

```
load_seeg(dataset = NULL, raw_data, geo_level, language = "eng")
```

Arguments

| | |
|-----------|--|
| dataset | A dataset name ("seeg", "seeg_farming", "seeg_industry", "seeg_energy", "seeg_land", "seeg_residuals"). On which "seeg" contains all five sectors (only works with <code>raw_data = TRUE</code>) and the others are filtered specifically by a main source of emission. See http://seeg.eco.br/notas-metodologicas/ |
| raw_data | A boolean setting the return of raw (TRUE) or processed (FALSE) data. |
| geo_level | A string that defines the geographic level of the data. Can be one of "country", "state" or "municipality". |
| language | A string that indicates in which language the data will be returned. Currently, only Portuguese ("pt") and English ("eng") are supported. Defaults to "eng". |

Value

A tibble with the selected data.

Examples

```
## Not run:  
# download state raw data  
seeg <- load_seeg(  
  dataset = "seeg",  
  raw_data = TRUE,  
  geo_level = "state"  
)  
  
# download country energy treated data  
seeg1 <- load_seeg(  
  dataset = "seeg_energy",  
  raw_data = FALSE,  
  geo_level = "country"  
)  
  
# download land municipality treated data
```

```
seeg2 <- load_seeg(  
  dataset = "seeg_land",  
  raw_data = FALSE,  
  geo_level = "municipality"  
)  
  
## End(Not run)
```

load_sigmine

SIGMINE - Mining Geographic Information System

Description

Loads information the mines being explored legally in Brazil, including their location, status, product being mined and area in square meters etc. Survey is done at municipal and state level

Usage

```
load_sigmine(dataset = "sigmine_active", raw_data, language = "eng")
```

Arguments

| | |
|----------|--|
| dataset | A dataset name ("sigmine_active") |
| raw_data | A boolean setting the return of raw (TRUE) or processed (FALSE) data. |
| language | A string that indicates in which language the data will be returned. Currently, only Portuguese ("pt") and English ("eng") are supported. Defaults to "eng". |

Value

A tibble with the selected data.

Examples

```
## Not run:  
# download state raw data  
sigmine_active <- load_sigmine(  
  dataset = "sigmine_active",  
  raw_data = TRUE  
)  
  
## End(Not run)
```

Index

load_baci, 2
load_br_trade, 3
load_cempre, 4
load_cipo, 5
load_climate, 6
load_degrad, 7
load_deter, 8
load_ibama, 9
load_ips, 10
load_mapbiomas, 11
load_pam, 12
load_pevs, 13
load_pibmunic, 14
load_ppm, 15
load_prodes, 16
load_seeg, 17
load_sigmine, 18