Package: psData (via r-universe)

January 30, 2025

Type Package

Title Download Regularly Maintained Political Science Data Sets

Version 0.2.2.9000

Date 2021-11-14

Author Christopher Gandrud

Maintainer Christopher Gandrud <christopher.gandrud@gmail.com>

Description This R package includes functions for gathering commonly
used and regularly maintained data set in political science. It
also includes functions for combining components from these
data sets into variables that have been suggested in the
literature, but are not regularly maintained.

License GPL (>= 3)

URL https://CRAN.R-project.org/package=psData,
https://ropengov.github.io/psData/,
https://github.com/rOpenGov/psData

BugReports https://github.com/rOpenGov/psData/issues

Depends R (>= 3.5.0)

Imports countrycode, DataCombine, readxl, reshape2, rio

Suggests knitr

Encoding UTF-8

LazyData TRUE

RoxygenNote 7.1.2

X-schema.org-isPartOf http://ropengov.org/

X-schema.org-keywords ropengov

Config/pak/sysreqs make libicu-dev libssl-dev libx11-dev zlib1g-dev

Repository https://ropengov.r-universe.dev

RemoteUrl https://github.com/rOpenGov/psData

RemoteRef HEAD

RemoteSha 13194238da1499975f1b8bb0f7eccd579b7d5546

Contents

CountryID   *Function for creating standardised country names and ID variables*

Description

Function for creating standardised country names and ID variables based on capabilities from the
countrycode package. The function also reports if duplicated country IDs have been created and
lets the user either drop these or return only duplicated values for inspection.

Usage

```
CountryID(
  data,
  countryVar = "country",
  OutCountryID = "iso2c",
  standardCountryName = TRUE,
  duplicates = "message",
  timeVar = NULL,
  fromLast = FALSE
)
```

Arguments

data	a data frame object
countryVar	character string naming the country.name variable. See countrycode.
OutCountryID	character string. The type of country ID you would like to include in the output file along with the country name. See countrycode for available options.
standardCountryName	
	logical. Whether or not to standardise the country names variable based on country.name from countrycode.
duplicates	character string specifying how to handle duplicated country or country-time observations (for the latter see timeVar). Can be set to none to do nothing, message to simply report duplicates, drop to report and drop duplicates, and return to return a data frame with only duplicated observations (see also fromLast).

timeVar	character string indicating the name of a time variable. For example, country time series often have separate rows based on a `year` variable. This is used solely to determine if there are duplicated country-time values.
fromLast	logical indicating if duplication should be considered from the reverse side. Only relevant if `duplicates = 'drop'` or `duplicates = 'return'`.

See Also

[duplicated](#)

DDGet	*Downloads the Democracy and Dictatorship data set*

Description

Downloads the Democracy and Dictatorship data set. It keeps specified variables and creates a standard country ID variable that can be used for merging the data with other data sets. See the codebook at the authors' website https://sites.google.com/site/joseantoniocheibub/datasets/dd?authuser=0

Usage

```
DDGet(
  url = "http://uofi.box.com/shared/static/bba3968d7c3397c024ec.dta",
  vars = NULL,
  OutCountryID = "iso2c",
  standardCountryName = TRUE,
  na.rm = TRUE,
  duplicates = "message",
  fromLast = FALSE
)
```

Arguments

url	character string. The URL for the Democracy and Dictatorship data set you would like to download. Note: it must be for the Stata version of the file.
vars	character vector containing the variables to keep. If `vars = NULL` then the entire data set is returned. Note that the `country` and `year` variables are always returned.
OutCountryID	character string. The type of country ID you would like to include in the output file along with the country name. See [countrycode](#) for available options.
standardCountryName	
	logical. Whether or not to standardise the country names variable based on `country.name` from [countrycode](#).
na.rm	logical. Drop observations where `OutCountryID` is NA.

| duplicates | character specifying how to handle duplicated country-year observations. Can be set to none to do nothing, `message` to simply report duplicates, `drop` to report and drop duplicates, and `return` to return a data frame with only duplicated observations (see also `fromLast`). |
| fromLast | logical indicating if duplication should be considered from the reverse side. Only relevant if `duplicates = 'drop'` or `duplicates = 'out'`. |

Value

a data frame

See Also

[countrycode](#), [CountryID](#)

Examples

```
## Not run:
# Download full data set
DDData <- DDGet()

## End(Not run)
```

DpiGet   *Downloads the Database of Political Institutions (DPI)*

Description

Downloads the Database of Political Institutions (DPI) data set. It keeps specified variables and creates a standard country ID variable that can be used for merging the data with other data sets.

Usage

```
DpiGet(
  url,
  vars = NULL,
  OutCountryID = "iso2c",
  standardCountryName = TRUE,
  na.rm = TRUE,
  duplicates = "message",
  fromLast = FALSE
)
```

Arguments

url	character string. The URL for the Polity IV data set you would like to download. Note this is exclusively to download previous, IMF hosted, versions of the data set. If a value is not supplied, then the 2015 IDB hosted version will be downloaded. If a link is supplied it must be to a Stata formated file.
vars	character vector containing the variables to keep. If vars = NULL then the entire data set is returned. Note that country and year variables are always returned.
OutCountryID	character string. The type of country ID you would like to include in the output file along with the country name. See [countrycode](#) for available options.
standardCountryName	
	logical. Whether or not to standardise the country names variable based on country.name from [countrycode](#).
na.rm	logical. Drop observations where OutCountryID is NA.
duplicates	character specifying how to handle duplicated country-year observations. Can be set to none to do nothing, message to simply report duplicates, drop to report and drop duplicates, and return to return a data frame with only duplicated observations (see also fromLast).
fromLast	logical indicating if duplication should be considered from the reverse side. Only relevant if duplicates = 'drop' or duplicates = 'out'.

Value

a data frame

See Also

[countrycode](#), [CountryID](#), [WinsetCreator](#)

Examples

```
## Not run:
# Download full data set
DpiData <- DpiGet()

# Create data frame with only the military variable
DpiSub <- DpiGet(vars = 'military',
                 OutCountryID = 'imf')

## End(Not run)
```

DropNA.psData	*Drop rows from a data frame with missing values in the OutCountryID variable.*

Description

Drop rows from a data frame with missing values in the OutCountryID variable.

Usage

```
DropNA.psData(data, countryVar = "country", timeVar = NULL, OutCountryID)
```

Arguments

data	a data frame object.
countryVar	a character vector naming the variables you would like to have only non-missing (NA) values.
timeVar	variable indicating the time.
OutCountryID	the ID the output country identifier.

Source

Largely based on `DropNA` from the `DataCombine` package.

IMF_WBGet	*Downloads Dreher's data set of IMF programs and World Bank projects (1970-2011)*

Description

Downloads Dreher's data set of IMF programs and World Bank projects (1970-2011). It keeps specified variables and creates a standard country ID variable that can be used for merging the data with other data sets.

Usage

```
IMF_WBGet(
  url = "http://axel-dreher.de/Dreher%20IMF%20and%20WB.xls",
  sheets = c("WB other agreed", "WB environment agreed"),
  OutCountryID = "iso2c",
  message = TRUE,
  standardCountryName = TRUE
)
```

Arguments

url	character string. The URL for the Dreher data set you would like to download. Note: it must be for the xlx version of the file. Currently only the 1970-2011 version is supported.
sheets	character vector of the Excel sheets (variables) that you would like to return. See Details for more information.
OutCountryID	character string. The type of country ID you would like to include in the output file along with the country name. See countrycode for available options.
message	logical. Whether or not to notify you which of sheets are being cleaned and organised.
standardCountryName	
	logical. Whether or not to standardise the country names variable based on country.name from countrycode.

Details

Using the sheets argument you can select which variables to download from their individual workbook seets in the original data set. These include:

- WB other agreed: Number of World Bank projects agreed, other than technical or adjustment.

- WB technical agreed: Number of World Bank technical projects agreed.

- WB adjustment agreed: Number of World Bank adjustment projects agreed.

- WB environment agreed: Number of World Bank environmental projects agreed.

- WB adjustment 5: Number of World Bank adjustment projects in effect for at least 5 months in a particular year.

- IMF SBA: IMF Standby Arrangement agreed, dummy.

- IMF EFF: IMF Extended Fund Facility Arrangement agreed, dummy.

- IMF SAF: IMF Structural Adjustment Facility Arrangement agreed, dummy.

- IMF PRGF: IMF Poverty Reduction and Growth Facility Arrangement agreed, dummy.

- IMF SBA 5: IMF Standby Arrangement in effect for at least 5 months in a particular year, dummy.

- IMF EFF 5: IMF Extended Fund Facility Arrangement in effect for at least 5 months in a particular year, dummy.

- IMF SAF 5: IMF Structural Adjustment Facility Arrangement in effect for at least 5 months in a particular year, dummy.

- IMF PRGF 5: IMF Poverty Reduction and Growth Facility Arrangement in effect for at least 5 months in a particular year, dummy.

Value

a data frame

Source

Data website: [http://www.uni-heidelberg.de/fakultaeten/wiso/awi/professuren/intwipol/](http://www.uni-heidelberg.de/fakultaeten/wiso/awi/professuren/intwipol/datasets_en.html)
[datasets_en.html](http://www.uni-heidelberg.de/fakultaeten/wiso/awi/professuren/intwipol/datasets_en.html).

When using the IMF data, please cite:

Dreher, Axel, 2006, IMF and Economic Growth: The Effects of Programs, Loans, and Compliance
with Conditionality, World Development 34, 5: 769-788.

When using the World Bank data, please cite:

Boockmann, Bernhard and Axel Dreher, 2003, The Contribution of the IMF and the World Bank to
Economic Freedom, European Journal of Political Economy 19, 3: 633-649.

Examples

```
## Not run:
# Download 'WB other agreed', 'WB environment agreed'
# These are the default sheets to gather
WBPrograms <- IMF_WBGet()

## End(Not run)
```

PolityGet   *Downloads Polity IV*

Description

Downloads the Polity IV data set. It keeps specified variables and creates a standard country ID
variable that can be used for merging the data with other data sets.

Usage

```
PolityGet(
  url = "http://www.systemicpeace.org/inscr/p4v2015.sav",
  vars = NULL,
  OutCountryID = "iso2c",
  standardCountryName = TRUE,
  na.rm = TRUE,
  duplicates = "message",
  fromLast = FALSE
)
```

Arguments

url   character string. The URL for the Polity IV data set you would like to download.
                   Note: it must be for the SPSS version of the file.

vars	character vector containing the variables to keep. If vars = NULL then the entire data set is returned. Note that the country and year variables are always returned.
OutCountryID	character string. The type of country ID you would like to include in the output file along with the country name. See countrycode for available options.
standardCountryName	
	logical. Whether or not to standardise the country names variable based on country.name from countrycode.
na.rm	logical. Drop observations where OutCountryID is NA.
duplicates	character specifying how to handle duplicated country-year observations. Can be set to none to do nothing, message to simply report duplicates, drop to report and drop duplicates, and return to return a data frame with only duplicated observations (see also fromLast).
fromLast	logical indicating if duplication should be considered from the reverse side. Only relevant if duplicates = 'drop' or duplicates = 'out'.

Value

a data frame

See Also

countrycode, CountryID, WinsetCreator

Examples

```
## Not run:
# Download full data set
PolityData <- PolityGet()

# Create data frame with only the main Polity democracy variable (polity2)
Polity2Data <- PolityGet(vars = 'polity2',
                         OutCountryID = 'imf')

## End(Not run)
```

RRCrisisGet	*Download and combine Reinhart and Rogoff's (2010) crisis dummy variables into one data frame*

Description

Download and combine Reinhart and Rogoff's (2010) crisis dummy variables into one data frame

Usage

```
RRCrisisGet(
  urls = c("http://www.carmenreinhart.com/user_uploads/data/22_data.xls",
    "http://www.carmenreinhart.com/user_uploads/data/35_data.xls",
    "http://www.carmenreinhart.com/user_uploads/data/23_data.xls",
    "http://www.carmenreinhart.com/user_uploads/data/25_data.xls"),
  OutCountryID = "iso2c",
  message = TRUE,
  standardCountryName = TRUE
)
```

Arguments

urls   URLs for each Excel file in the Reinhart and Rogoff data set. See [https://](https://carmenreinhart.com/browse-by-topic/)
                carmenreinhart.com/browse-by-topic/.

OutCountryID   character string. The type of country ID you would like to include in the output
                file along with the country name. See countrycode for available options.

message   logical. Whether or not to notify you which of sheets are being cleaned and
                organised.

standardCountryName

                logical. Whether or not to standardise the country names variable based on
                country.name from countrycode.

Value

Returns a data frame with the following columns:

- iso2c: The ISO two letter country code identifying the country. This can be changed to
  another country ID system using OutCountryID

- country: Country names.

- year: The year.

- RR_Independence: Year of independence.

- RR_CurrencyCrisis: Currency crisis.

- RR_InflationCrisis: Inflation crisis.

- RR_StockMarketCrash: Stock market crash.

- RR_SovDebtCrisisDom: Domestic sovereign debt crisis.

- RR_SovDebtCrisisExt: External sovereign debt crisis.

- RR_BankingCrisis: Banking crisis.

- RR_YearlyCrisisTally: Total number of crises per year.

Source

Reinhart, Camen M. and Kenneth S. Rogoff, "From Financial Crash to Debt Crisis," NBER Working
Paper 15795, March 2010. Forthcoming in American Economic Review.

Examples

```
## Not run:
 RRDummies <- RRCrisisGet()

## End(Not run)
```

WinsetCreator   *Create Winset and Selectorate variables*

Description

Creates the winset (W) and a modified version of the selectorate (S) variable from Bueno de Mesquita et al. (2003) using the most recent data available from Polity IV and the Database of Political Institutions.

Usage

```
WinsetCreator(
  PolityUrl = "http://www.systemicpeace.org/inscr/p4v2015.sav",
  DpiUrl = NULL,
  OutCountryID = "iso2c",
  na.rm = TRUE
)
```

Arguments

PolityUrl	character string. The URL for the Polity IV data set you would like to download. Note: it must be for the SPSS version of the file.
DpiUrl	character string. The URL for the Database of Political Institutions data set you would like to download. Note: the link must be to a Stata formated file.
OutCountryID	character string. The type of country ID you would like to include in the output file along with the country name. See [countrycode](#) for available options.
na.rm	logical. Drop observations where OutCountryID is NA.

Value

Returns a data frame with the following columns:

- iso2c: The ISO two letter country code identifying the country. This can be changed to another country ID system using OutCountryID

- country: Country names.

- year: The year.

- W: The winset variable. The variable is the same as Bueno de Mesquita (2003) except the military executive component is from DPI, rather than Banks (1996).

- `ModS`: The modified selectorate variable. Instead of being based on Polity LEGSELEC, which is no longer create, `ModS` is based on the Legislative Index of Electoral Competetiveness variable (LIEC) from DPI. No legislature = 1. Unelected legislature = 2. Elected legislature with only 1 candidate = 3. Elected legislatures with 1 party and multiple candidates = 4. All other legislatures = 5. The variable is then standardised between 0 and 1 by subtracting by 1 then dividing by 4. Note: some countries were coded in LIEC as 3.5, so `ModS` has 6 levels between 0 and 1.

Source

See Bueno de Mesquita Bruce, Alastair Smith, Randolph M. Siverson, and James D. Morrow. 2003. The Logic of Political Survival. Cambridge, MA: MIT Press.

Morrow, J. D., Bueno De Mesquita, B., Siverson, R. M., and Smith, A. 2008. Retesting Selectorate Theory: Separating the Effects of W from Other Elements of Democracy. American Political Science Review, 102(03), 393-400.

Polity IV: http://www.systemicpeace.org/.

Modified so that military regime is take from Database of Political Institutions (https://publications.iadb.org/en/database-political-institutions-2020-dpi2020), rather than Banks, Arthur S. 1996. Political Handbook of the World. New York: CSA Publications.

See Also

countrycode, CountryID, DpiGet, PolityGet

Examples

```
## Not run:
# Create winset data using default options
WinsetData <- WinsetCreator()

## End(Not run)
```

Index