

# Package ‘wbstats’

January 3, 2018

**Type** Package

**Title** Programmatic Access to Data and Statistics from the World Bank API

**Description** Tools for searching and downloading data and statistics from the World Bank Data API (<<http://data.worldbank.org/developers/api-overview>>) and the World Bank Data Catalog API (<<http://data.worldbank.org/developers/data-catalog-api>>).

**Version** 0.2

**Date** 2018-01-01

**Maintainer** Jesse Piburn <[piburnjo@ornl.gov](mailto:piburnjo@ornl.gov)>

**URL** <https://github.com/GIST-ORNL/wbstats>

**License** MIT + file LICENSE

**LazyData** TRUE

**Depends** R (>= 3.1)

**Imports** httr, jsonlite, tidyr

**Suggests** ggplot2, knitr, lubridate (>= 1.5.0)

**VignetteBuilder** knitr

**RoxygenNote** 6.0.1

**NeedsCompilation** no

**Author** Jesse Piburn [aut, cre],  
UT-Battelle, LLC [cph]

**Repository** CRAN

**Date/Publication** 2018-01-03 18:21:36 UTC

## R topics documented:

|                         |   |
|-------------------------|---|
| call_api . . . . .      | 2 |
| wb . . . . .            | 3 |
| wbcache . . . . .       | 5 |
| wbcountries . . . . .   | 7 |
| wbdatacatalog . . . . . | 8 |

|                          |    |
|--------------------------|----|
| wbdate2POSIXct . . . . . | 8  |
| wbformatcols . . . . .   | 9  |
| wbget . . . . .          | 9  |
| wbget_dc . . . . .       | 10 |
| wbincome . . . . .       | 10 |
| wbindicators . . . . .   | 11 |
| wblending . . . . .      | 12 |
| wbsearch . . . . .       | 13 |
| wbsources . . . . .      | 13 |
| wbstats . . . . .        | 14 |
| wbtopics . . . . .       | 15 |
| wburls . . . . .         | 16 |
| wb_cachelist . . . . .   | 16 |

|              |           |
|--------------|-----------|
| <b>Index</b> | <b>17</b> |
|--------------|-----------|

---

|          |  |
|----------|--|
| call_api | <i>Call the World Bank API and return list</i> |
|----------|--|

---

### Description

To be used inside of wbget()

### Usage

```
call_api(url_string, indicator)
```

### Arguments

|            |   |
|------------|---|
| url_string | A character string. A formatted url string                          |
| indicator  | A character string. indicatorID for request. Used for error returns |

### Value

json contents of page information

---

 wb *Download Data from the World Bank API*


---

**Description**

This function downloads the requested information using the World Bank API

**Usage**

```
wb(country = "all", indicator, startdate, enddate, mrv, return_wide = FALSE,
    gapfill, freq, cache, lang = c("en", "es", "fr", "ar", "zh"),
    removeNA = TRUE, POSIXct = FALSE, include_dec = FALSE,
    include_unit = FALSE, include_obsStatus = FALSE,
    include_lastUpdated = FALSE)
```

**Arguments**

|             |  |
|-------------|--|
| country     | Character vector of country or region codes. Default value is special code of all. Other permissible values are codes in the following fields from the <a href="#">wb_cachelist</a> country data frame. iso3c, iso2c, regionID, adminID, and incomeID. Additional special values include aggregates, which returns only aggregates, and countries_only, which returns all countries without aggregates.                      |
| indicator   | Character vector of indicator codes. These codes correspond to the indicatorID column from the indicator data frame of <a href="#">wbcache</a> or <a href="#">wb_cachelist</a> , or the result of <a href="#">wbindicators</a>   |
| startdate   | Numeric or character. If numeric it must be in %Y form (i.e. four digit year). For data at the subannual granularity the API supports a format as follows: for monthly data, "2016M01" and for quarterly data, "2016Q1". This also accepts a special value of "YTD", useful for more frequently updated subannual indicators.  |
| enddate     | Numeric or character. If numeric it must be in %Y form (i.e. four digit year). For data at the subannual granularity the API supports a format as follows: for monthly data, "2016M01" and for quarterly data, "2016Q1".   |
| mrv         | Numeric. The number of Most Recent Values to return. A replacement of startdate and enddate, this number represents the number of observations you which to return starting from the most recent date of collection. Useful in conjunction with freq   |
| return_wide | Logical. If TRUE data is returned in a wide format instead of long, with a column named for each indicatorID. To necessitate this transformation, the indicator column, that provides the human readable description is dropped. This field is available through from the indicator data frame of <a href="#">wbcache</a> or <a href="#">wb_cachelist</a> , or the result of <a href="#">wbindicators</a> . Default is FALSE |
| gapfill     | Logical. Works with mrv. If TRUE fills values, if not available, by back tracking to the next available period (max number of periods back tracked will be limited by mrv number)  |

|                     |   |
|---------------------|---|
| freq                | Character String. For fetching quarterly ("Q"), monthly("M") or yearly ("Y") values. Currently works along with <code>mr.v</code> . Useful for querying high frequency data.  |
| cache               | List of data frames returned from <code>wbcache</code> . If omitted, <code>wb_cachelist</code> is used  |
| lang                | Language in which to return the results. If lang is unspecified, english is the default.  |
| removeNA            | if TRUE, remove any blank or NA observations that are returned. if FALSE, no blank or NA values are removed from the return.  |
| POSIXct             | if TRUE, additional columns <code>date_ct</code> and <code>granularity</code> are added. <code>date_ct</code> converts the default date into a <code>POSIXct</code> . <code>granularity</code> denotes the time resolution that the date represents. Useful for subannual data and mixing subannual with annual data. If FALSE, these fields are not added. |
| include_dec         | if TRUE, the column decimal is not removed from the return. if FALSE, this column is removed  |
| include_unit        | if TRUE, the column unit is not removed from the return. if FALSE, this column is removed   |
| include_obsStatus   | if TRUE, the column <code>obsStatus</code> is not removed from the return. if FALSE, this column is removed   |
| include_lastUpdated | if TRUE, the column <code>lastUpdated</code> is not removed from the return. if FALSE, this column is removed. If TRUE and <code>POSIXct = TRUE</code> then column will be of class <code>Date</code>   |

### Value

Data frame with all available requested data.

### Note

Not all data returns have support for languages other than english. If the specific return does not support your requested language by default it will return NA. For an enumeration of supported languages by data source please see [wbdatacatalog](#). The options for lang are:

- en: English
- es: Spanish
- fr: French
- ar: Arabic
- zh: Mandarin

The `POSIXct` parameter requires the use of `lubridate` ( $\geq 1.5.0$ ). All dates are rounded down to the floor. For example a value for the year 2016 would have a `POSIXct` date of `2016-01-01`. If this package is not available and the `POSIXct` parameter is set to TRUE, the parameter is ignored and a warning is produced.

The `include_dec`, `include_unit`, and `include_obsStatus` are defaulted to FALSE because as of writing, all returns have a value of  $\emptyset$ , NA, and NA, respectively. These columns might be used in the future by the API, therefore the option to include the column is available.

The `include_lastUpdated` is defaulted to `FALSE` as well to limit the  
 If there is no data available that matches the request parameters, an empty data frame is returned along with a warning. This design is for easy aggregation of multiple calls.

### Examples

```
# GDP at market prices (current US$) for all available countries and regions
wb(indicator = "NY.GDP.MKTP.CD", startdate = 2000, enddate = 2016)

# GDP and Population in long format for the most recent 20 observations
wb(indicator = c("SP.POP.TOTL", "NY.GDP.MKTP.CD"), mrv = 20)

# GDP and Population in wide format for the most recent 20 observations
wb(indicator = c("SP.POP.TOTL", "NY.GDP.MKTP.CD"), mrv = 20, return_wide = TRUE)

# query using regionID or incomeID
# High Income Countries and Sub-Saharan Africa (all income levels)
wb(country = c("HIC", "SSF"), indicator = "NY.GDP.MKTP.CD", startdate = 1985, enddate = 1985)

# if you do not know when the latest time an indicator is available mrv can help
wb(country = c("IN"), indicator = 'EG.ELC.ACCS.ZS', mrv = 1)

# increase the mrv value to increase the number of maximum number of returns
wb(country = c("IN"), indicator = 'EG.ELC.ACCS.ZS', mrv = 35)

# GDP at market prices (current US$) for only available countries
wb(country = "countries_only", indicator = "NY.GDP.MKTP.CD", startdate = 2000, enddate = 2016)

# GDP at market prices (current US$) for only available aggregate regions
wb(country = "aggregates", indicator = "NY.GDP.MKTP.CD", startdate = 2000, enddate = 2016)

# if you want to "fill-in" the values in between actual observations use gapfill = TRUE
# this highlights a very important difference.
# all other parameters are the same as above, except gapfill = TRUE
# and the results are very different
wb(country = c("IN"), indicator = 'EG.ELC.ACCS.ZS', mrv = 35, gapfill = TRUE)

# if you want the most recent values within a certain time frame
wb(country = c("US"), indicator = 'SI.DST.04TH.20', startdate = 1970, enddate = 2000, mrv = 2)

# without the freq parameter the default temporal granularity search is yearly
# should return the 12 most recent years of data
wb(country = c("CHN", "IND"), indicator = "DPANUSSPF", mrv = 12)

# if another frequency is available for that indicator it can be accessed using the freq parameter
# should return the 12 most recent months of data
wb(country = c("CHN", "IND"), indicator = "DPANUSSPF", mrv = 12, freq = "M")
```

## Description

Download an updated list of information regarding countries, indicators, sources, data catalog, indicator topics, lending types, and income levels from the World Bank API

## Usage

```
wbcache(lang = c("en", "es", "fr", "ar", "zh"))
```

## Arguments

`lang` Language in which to return the results. If `lang` is unspecified, english is the default.

## Value

A list containing the following items:

- `countries`: A data frame. The result of calling [wbcountries](#)
- `indicators`: A data frame. The result of calling [wbindicators](#)
- `sources`: A data frame. The result of calling [wbsources](#)
- `datacatalog`: A data frame. The result of calling [wbdatacatalog](#)
- `topics`: A data frame. The result of calling [wbtopics](#)
- `income`: A data frame. The result of calling [wbincome](#)
- `lending`: A data frame. The result of calling [wblending](#)

## Note

Not all data returns have support for languages other than english. If the specific return does not support your requested language by default it will return NA. For an enumeration of supported languages by data source please see [wbdatacatalog](#). The options for `lang` are:

- `en`: English
- `es`: Spanish
- `fr`: French
- `ar`: Arabic
- `zh`: Mandarin

List item `datacatalog` will always return in english, as the API does not support any other languages for that information.

Saving this return and using it has the `cache` parameter in [wb](#) and [wbsearch](#) replaces the default cached version [wb\\_cachelist](#) that comes with the package itself

## Examples

```
# default is english. To specific another language use argument lang
wbcache(lang = "es")
```

---

|             |  |
|-------------|--|
| wbcountries | <i>Download updated country and region information from World Bank API</i> |
|-------------|--|

---

### Description

Download updated information on available countries and regions from the World Bank API

### Usage

```
wbcountries(lang = c("en", "es", "fr", "ar", "zh"))
```

### Arguments

|      |  |
|------|--|
| lang | Language in which to return the results. If lang is unspecified, english is the default. |
|------|--|

### Value

A data frame of available countries and regions with related information

### Note

Not all data returns have support for languages other than english. If the specific return does not support your requested language by default it will return NA. For an enumeration of supported languages by data source please see [wbdatacatalog](#). The options for lang are:

- en: English
- es: Spanish
- fr: French
- ar: Arabic
- zh: Mandarin

### Examples

```
wbcountries()
```

```
wbcountries(lang = "es")
```

---

|               |  |
|---------------|--|
| wbdatacatalog | <i>Download an updated list of the World Bank data catalog</i> |
|---------------|--|

---

**Description**

Download an updated list of the World Bank data catalog from the World Bank API

**Usage**

```
wbdatacatalog()
```

**Value**

A data frame of the World Bank data catalog with related information

**Note**

This function does not support any languages other than english due to the lack of support from the World Bank API

**Examples**

```
wbdatacatalog()
```

---

|                |   |
|----------------|---|
| wbdate2POSIXct | <i>Add a POSIXct dates to a World Bank API return</i> |
|----------------|---|

---

**Description**

Add a POSIXct date column as well as a column with the appropriate granularity to a World Bank API return

**Usage**

```
wbdate2POSIXct(df, date_col)
```

**Arguments**

|          |                                   |
|----------|-----------------------------------|
| df       | data frame returned from API call |
| date_col | name of the current date field    |

**Value**

If the package lubridate ( $\geq 1.5.0$ ) is available the original data frame with two new columns, `data_ct` and `granularity` is returned. If the above package is not available, the original data frame is returned unaltered with an additional warning message.



---

|              |  |
|--------------|--|
| wbformatcols | <i>Format column names of World Bank API returns</i> |
|--------------|--|

---

**Description**

change the name of columns that we know of to our standardized names. If there are new columns added that we don't know about, leave their name the same

**Usage**

```
wbformatcols(df, col_names, blank2NA = TRUE)
```

**Arguments**

|           |   |
|-----------|---|
| df        | data frame returned from API call                               |
| col_names | a named vector of column names and what they will be changed to |
| blank2NA  | logical. If TRUE blank values are converted to NA               |

**Value**

a data frame with the column names changed accordingly

---

|       |  |
|-------|--|
| wbget | <i>Call the World Bank API and return a formatted data frame</i> |
|-------|--|

---

**Description**

This function calls `wbget.raw` and determines if there are multiple pages from the request. If there are multiple pages then it generates a list of data frames from each page and then combines the results with `do.call("rbind", mylist)`

**Usage**

```
wbget(url_string, indicator)
```

**Arguments**

|            |   |
|------------|---|
| url_string | A character string. A formatted url string                          |
| indicator  | A character string. indicatorID for request. Used for error returns |

**Value**

A data frame

---

`wbget_dc`*Call the Data Catalog API*

---

**Description**

Helper function for the data catalog call

**Usage**

```
wbget_dc(url_string)
```

**Arguments**

`url_string` A character string. A formatted url string

**Value**

A list of data frames

**Note**

This call is separate because the data catalog is actually a different API and therefore has a different return structure.

---

`wbincome`*Download updated income type information from World Bank API*

---

**Description**

Download updated information on available income types from the World Bank API

**Usage**

```
wbincome(lang = c("en", "es", "fr", "ar", "zh"))
```

**Arguments**

`lang` Language in which to return the results. If `lang` is unspecified, english is the default.

**Value**

A data frame of available income types with related information

**Note**

Not all data returns have support for languages other than english. If the specific return does not support your requested language by default it will return NA. For an enumeration of supported languages by data source please see [wdatacatalog](#). The options for lang are:

- en: English
- es: Spanish
- fr: French
- ar: Arabic
- zh: Mandarin

**Examples**

```
wbincome()
```

```
wbincome(lang = "es")
```

---

wbindicators

*Download updated indicator information from World Bank API*

---

**Description**

Download updated information on available indicators from the World Bank API

**Usage**

```
wbindicators(lang = c("en", "es", "fr", "ar", "zh"))
```

**Arguments**

|      |  |
|------|--|
| lang | Language in which to return the results. If lang is unspecified, english is the default. |
|------|--|

**Value**

A data frame of available indicators with related information

**Note**

Not all data returns have support for languages other than english. If the specific return does not support your requested language by default it will return NA. For an enumeration of supported languages by data source please see [wdatacatalog](#). The options for lang are:

- en: English
- es: Spanish
- fr: French
- ar: Arabic
- zh: Mandarin

## Examples

```
# default is english. To specific another language use argument lang
wbindicators(lang = "es")
```

---

wblending

*Download updated lending type information from World Bank API*

---

## Description

Download updated information on available lending types from the World Bank API

## Usage

```
wblending(lang = c("en", "es", "fr", "ar", "zh"))
```

## Arguments

lang            Language in which to return the results. If lang is unspecified, english is the default.

## Value

A data frame of available lending types with related information

## Note

Not all data returns have support for languages other than english. If the specific return does not support your requested language by default it will return NA. For an enumeration of supported languages by data source please see [wbdatacatalog](#). The options for lang are:

- en: English
- es: Spanish
- fr: French
- ar: Arabic
- zh: Mandarin

## Examples

```
wblending()
```

```
wblending(lang = "es")
```

---

|          |  |
|----------|--|
| wbsearch | <i>Search indicator information available through the World Bank API</i> |
|----------|--|

---

**Description**

This function allows finds indicators that match a search term and returns a data frame of matching results

**Usage**

```
wbsearch(pattern = "poverty", fields = c("indicator", "indicatorDesc"),
  extra = FALSE, cache)
```

**Arguments**

|         |  |
|---------|--|
| pattern | Character string or regular expression to be matched   |
| fields  | Character vector of column names through which to search   |
| extra   | if FALSE, only the indicator ID and short name are returned, if TRUE, all columns of the cache parameter's indicator data frame are returned |
| cache   | List of data frames returned from <a href="#">wbcache</a> . If omitted, <a href="#">wb_cachelist</a> is used                                 |

**Value**

Data frame with indicators that match the search pattern.

**Examples**

```
wbsearch(pattern = "education")

wbsearch(pattern = "Food and Agriculture Organization", fields = "sourceOrg")

# with regular expression operators
# 'poverty' OR 'unemployment' OR 'employment'
wbsearch(pattern = "poverty|unemployment|employment")
```

---

|           |   |
|-----------|---|
| wbsources | <i>Download updated data source information from World Bank API</i> |
|-----------|---|

---

**Description**

Download updated information on available data sources from the World Bank API

**Usage**

```
wbsources(lang = c("en", "es", "fr", "ar", "zh"))
```

## Arguments

`lang` Language in which to return the results. If `lang` is unspecified, english is the default.

## Value

A data frame of available data sources with related information

## Note

Not all data returns have support for languages other than english. If the specific return does not support your requested language by default it will return NA. For an enumeration of supported languages by data source please see [wbdatacatalog](#). The options for `lang` are:

- en: English
- es: Spanish
- fr: French
- ar: Arabic
- zh: Mandarin

## Examples

```
wbsources()  
wbsources(lang = "es")
```

---

|                      |  |
|----------------------|--|
| <code>wbstats</code> | <i>wbstats: An R package for searching and downloading data from the World Bank API.</i> |
|----------------------|--|

---

## Description

The `wbstats` package provides structured access to all of the information available from the World Bank API including; support for multiple languages, access to all annual, quarterly, and monthly data, as well as access to the World Bank Data Catalog API, which provides metadata on each data source, such as update schedules, last revisions, and coverage, among other information.

## Details

To learn more about the `wbstats` package, start with the vignettes: `browseVignettes(package = "wbstats")`

---

`wbtopics`*Download updated indicator topic information from World Bank API*

---

### Description

Download updated information on available indicator topics from the World Bank API

### Usage

```
wbtopics(lang = c("en", "es", "fr", "ar", "zh"))
```

### Arguments

|                   |   |
|-------------------|---|
| <code>lang</code> | Language in which to return the results. If <code>lang</code> is unspecified, english is the default. |
|-------------------|---|

### Value

A data frame of available indicator topics with related information

### Note

Not all data returns have support for languages other than english. If the specific return does not support your requested language by default it will return NA. For an enumeration of supported languages by data source please see [wbdatacatalog](#). The options for `lang` are:

- en: English
- es: Spanish
- fr: French
- ar: Arabic
- zh: Mandarin

### Examples

```
wbtopics()
```

```
wbtopics(lang = "es")
```

---

|        |   |
|--------|---|
| wburls | <i>url chunks to be used in API calls</i> |
|--------|---|

---

**Description**

This function is used inside other functions in this package

**Usage**

```
wburls()
```

**Value**

a list with a base url and a url section for formatting the json return

---

|              |   |
|--------------|---|
| wb_cachelist | <i>Cached information from the World Bank API</i> |
|--------------|---|

---

**Description**

This data is a cached result of the [wbcache](#) function. By default functions [wb](#) and [wbsearch](#) use this data for the cache parameter.

**Usage**

```
wb_cachelist
```

**Format**

A list containing 7 data frames:

- **countries**: A data frame. The result of calling [wbcountries](#)
- **indicators**: A data frame. The result of calling [wbindicators](#)
- **sources**: A data frame. The result of calling [wbsources](#)
- **datacatalog**: A data frame. The result of calling [wbdatacatalog](#)
- **topics**: A data frame. The result of calling [wbtopics](#)
- **income**: A data frame. The result of calling [wbincome](#)
- **lending**: A data frame. The result of calling [wblending](#)

**Details**

This data was updated using [wbcache](#) on January 1, 2018



# Index

## \*Topic **datasets**

- wb\_cachelist, [16](#)
- call\_api, [2](#)
- Date, [4](#)
- lubridate, [4](#)
- POSIXct, [4](#)
- wb, [3](#), [6](#), [16](#)
- wb\_cachelist, [3](#), [4](#), [6](#), [13](#), [16](#)
- wbcache, [3](#), [4](#), [5](#), [13](#), [16](#)
- wbcountries, [6](#), [7](#), [16](#)
- wbdatacatalog, [4](#), [6](#), [7](#), [8](#), [11](#), [12](#), [14–16](#)
- wbdate2POSIXct, [8](#)
- wbformatcols, [9](#)
- wbget, [9](#)
- wbget\_dc, [10](#)
- wbincome, [6](#), [10](#), [16](#)
- wbindicators, [3](#), [6](#), [11](#), [16](#)
- wblending, [6](#), [12](#), [16](#)
- wbsearch, [6](#), [13](#), [16](#)
- wbsources, [6](#), [13](#), [16](#)
- wbstats, [14](#)
- wbstats-package (wbstats), [14](#)
- wbtopics, [6](#), [15](#), [16](#)
- wburls, [16](#)