|  |  |
| --- | --- |
| cbs\_get\_data {cbsodataR} | R Documentation |

**Get data from Statistics Netherlands (CBS)**

**Description**

Retrieves data from a table of Statistics Netherlands. A list of available tables can be retrieved with [cbs\_get\_toc()](http://127.0.0.1:11903/library/cbsodataR/help/cbs_get_toc). Use the Identifier column of cbs\_get\_toc as id in cbs\_get\_data and cbs\_get\_meta.

**Usage**

cbs\_get\_data(

id,

...,

catalog = "CBS",

select = NULL,

typed = TRUE,

add\_column\_labels = TRUE,

dir = tempdir(),

verbose = FALSE,

base\_url = getOption("cbsodataR.base\_url", BASE\_URL),

include\_ID = FALSE

)

**Arguments**

|  |  |
| --- | --- |
| id | Identifier of table, can be found in [cbs\_get\_toc()](http://127.0.0.1:11903/library/cbsodataR/help/cbs_get_toc) |
| ... | optional filter statements, see details. |
| catalog | catalog id, can be retrieved with [cbs\_get\_datasets()](http://127.0.0.1:11903/library/cbsodataR/help/cbs_get_datasets) |
| select | character optional, columns to select |
| typed | Should the data automatically be converted into integer and numeric? |
| add\_column\_labels | Should column titles be added as a label (TRUE) which are visible in View |
| dir | Directory where the table should be downloaded. Defaults to temporary directory |
| verbose | Print extra messages what is happening. |
| base\_url | optionally specify a different server. Useful for third party data services implementing the same protocol. |
| include\_ID | Should the data include the ID column for the rows? |

**Details**

To reduce the download time, optionaly the data can be filtered on category values: for large tables (> 100k records) this is a wise thing to do.

The filter is specified with (see examples below):

* ⁠<column\_name> = <values>⁠ in which ⁠<values>⁠ is a character vector. Rows with values that are not part of the character vector are not returned. **Note that the values have to be values from the ⁠$Key⁠ column of the corresponding meta data. These may contain trailing spaces...**
* ⁠<column\_name> = has\_substring(x)⁠ in which x is a character vector. Rows with values that do not have a substring that is in x are not returned. Useful substrings are "JJ", "KW", "MM" for Periods (years, quarters, months) and "PV", "CR" and "GM" for Regions (provinces, corops, municipalities).
* ⁠<column\_name> = eq(<values>) | has\_substring(x)⁠, which combines the two statements above.

By default the columns will be converted to their type (typed=TRUE). CBS uses multiple types of missing (unknown, surpressed, not measured, missing): users wanting all these nuances can use typed=FALSE which results in character columns.

**Value**

data.frame with the requested data. Note that a csv copy of the data is stored in dir.

**Note**

All data are downloaded using [cbs\_download\_table()](http://127.0.0.1:11903/library/cbsodataR/help/cbs_download_table)

**See Also**

[cbs\_get\_meta()](http://127.0.0.1:11903/library/cbsodataR/help/cbs_get_meta), [cbs\_download\_data()](http://127.0.0.1:11903/library/cbsodataR/help/cbs_download_data)

Other data retrieval: [cbs\_add\_date\_column](http://127.0.0.1:11903/library/cbsodataR/help/cbs_add_date_column)(), [cbs\_add\_label\_columns](http://127.0.0.1:11903/library/cbsodataR/help/cbs_add_label_columns)(), [cbs\_download\_data](http://127.0.0.1:11903/library/cbsodataR/help/cbs_download_data)(), [cbs\_extract\_table\_id](http://127.0.0.1:11903/library/cbsodataR/help/cbs_extract_table_id)(), [cbs\_get\_data\_from\_link](http://127.0.0.1:11903/library/cbsodataR/help/cbs_get_data_from_link)()

Other query: [eq](http://127.0.0.1:11903/library/cbsodataR/help/eq)(), [has\_substring](http://127.0.0.1:11903/library/cbsodataR/help/has_substring)()

**Examples**

[Run examples](http://127.0.0.1:11903/library/cbsodataR/Example/cbs_get_data)

## Not run:

cbs\_get\_data( id = "7196ENG" # table id

, Periods = "2000MM03" # March 2000

, CPI = "000000" # Category code for total

)

# useful substrings:

## Periods: "JJ": years, "KW": quarters, "MM", months

## Regions: "NL", "PV": provinces, "GM": municipalities

cbs\_get\_data( id = "7196ENG" # table id

, Periods = has\_substring("JJ") # all years

, CPI = "000000" # Category code for total

)

cbs\_get\_data( id = "7196ENG" # table id

, Periods = c("2000MM03","2001MM12") # March 2000 and Dec 2001

, CPI = "000000" # Category code for total

)

# combine either this

cbs\_get\_data( id = "7196ENG" # table id

, Periods = has\_substring("JJ") | "2000MM01" # all years and Jan 2001

, CPI = "000000" # Category code for total

)

# or this: note the "eq" function

cbs\_get\_data( id = "7196ENG" # table id

, Periods = eq("2000MM01") | has\_substring("JJ") # Jan 2000 and all years

, CPI = "000000" # Category code for total

)

## End(Not run)

[Package *cbsodataR* version 0.5.1 [Index](http://127.0.0.1:11903/library/cbsodataR/html/00Index.html)]