

Package ‘cansim’

August 19, 2026

Type Package

Title Accessing Statistics Canada Data Table and Vectors

Version 0.5.0

Maintainer Jens von Bergmann <jens@mountainmath.ca>

Description Searches for, accesses, and retrieves Statistics Canada data tables, as well as individual vectors, as tidy data frames. This package enriches the tables with metadata, deals with encoding issues, allows for bilingual English or French language data retrieval, and bundles convenience functions to make it easier to work with retrieved table data. For more efficient data access the package allows for caching data in a local database and database level filtering, data manipulation and summarizing.

License MIT + file LICENSE

Encoding UTF-8

ByteCompile yes

NeedsCompilation no

LazyData true

Depends R (>= 4.1)

Imports digest (>= 0.6), dplyr (>= 1.1), httr2 (>= 1.1.0), tidyr (>= 1.3), readr (>= 2.1), rlang (>= 1.1), stringr (>= 1.5), purrr (>= 1.0), tibble (>= 3.2), arrow (>= 18.1), DBI (>= 1.2), RSQLite (>= 2.3), utils (>= 4.3), dbplyr (>= 2.5)

Suggests knitr, rmarkdown, ggplot2, scales, testthat (>= 3.0.0), withr

URL <https://github.com/mountainMath/cansim>,
<https://mountainmath.github.io/cansim/>,
<https://www.statcan.gc.ca/>

BugReports <https://github.com/mountainMath/cansim/issues>

VignetteBuilder knitr

Language en-CA

Config/testthat/edition 3

Config/roxygen2/version 8.0.0

Author Jens von Bergmann [aut, cre],
Dmitry Shkolnik [aut]

Repository CRAN

Date/Publication 2026-08-19 21:50:02 UTC

Contents

add_cansim_vectors_to_template	3
add_provincial_abbreviations	4
cansim_old_to_new	4
cansim_repartition_cached_table	5
categories_for_level	6
collect_and_normalize	7
disconnect_cansim_connection	8
get_cansim	9
get_cansim_changed_series_data_for_coordinates	10
get_cansim_changed_series_data_for_vectors	11
get_cansim_changed_series_list	12
get_cansim_changed_tables	13
get_cansim_code_set	14
get_cansim_column_categories	15
get_cansim_column_list	16
get_cansim_connection	17
get_cansim_cube_metadata	18
get_cansim_data_for_table_coord_periods	19
get_cansim_key_release_schedule	20
get_cansim_series_info_cube_coord	21
get_cansim_table_info	22
get_cansim_table_last_release_date	23
get_cansim_table_notes	23
get_cansim_table_overview	24
get_cansim_table_short_notes	25
get_cansim_table_subject	26
get_cansim_table_survey	27
get_cansim_table_template	28
get_cansim_table_url	29
get_cansim_vector	30
get_cansim_vector_for_latest_periods	31
get_cansim_vector_info	33
list_cansim_cached_tables	33
list_cansim_cubes	34
remove_cansim_cached_tables	35
search_cansim_cubes	36
set_cansim_cache_path	36
show_cansim_cache_path	37
view_cansim_webpage	37

`add_cansim_vectors_to_template`*Retrieve series info for given table id and coordinates*

Description

Retrieves vector information for given table and coordinates. This can be used to query data by vectors, it only returns vector information on coordinates which are present in the data table, so it gives an effective way to filter coordinates. Vector information is not available for census data tables.

Usage

```
add_cansim_vectors_to_template(template, refresh = FALSE)
```

Arguments

<code>template</code>	A (possibly filtered) cansim table template as returned by ‘get_cansim_table_template’
<code>refresh</code>	Refresh the data from the Statistics Canada API

Value

a tibble containing the table template with added vector information

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
## Not run:
template <- get_cansim_table_template("34-10-0013")
template |>
  filter(Geography=="Canada") |>
  add_cansim_vectors_to_template()

## End(Not run)
```

```
add_provincial_abbreviations
```

Add provincial abbreviations as factor

Description

Add provincial abbreviations as factor

Usage

```
add_provincial_abbreviations(data)
```

Arguments

`data` A tibble as returned by `get_cansim` with provincial level data

Value

The input tibble with additional factor `GEO.abb` that contains language-specific provincial abbreviations

Examples

```
## Not run:
df <- get_cansim("17-10-0005")
df <- add_provincial_abbreviations(df)

## End(Not run)
```

```
cansim_old_to_new
```

Translate deprecated CANSIM table number into new NDM-format table catalogue number

Description

Returns NDM table catalogue equivalent given a standard old-format CANSIM table number

Usage

```
cansim_old_to_new(oldCansimTableNumber)
```

Arguments

`oldCansimTableNumber`
 deprecated style CANSIM table number (e.g. "427-0001")

Value

A character string with the new-format NDM table number

Examples

```
cansim_old_to_new("026-0018")
```

```
cansim_repartition_cached_table
```

Repartitions a cached cansim table to a new partitioning scheme

Description

Repartitions and already downloaded and cached parquet or feather dataset

Usage

```
cansim_repartition_cached_table(
  cansimTableNumber,
  new_partitioning = c(),
  language = "english",
  format = "parquet",
  cache_path = Sys.getenv("CANSIM_CACHE_PATH")
)
```

Arguments

cansimTableNumber	the NDM table number to load
new_partitioning	(Optional) Partition columns to use for parquet or feather formats.
language	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored
format	(Optional) The format of the data table to retrieve. Either "parquet", "feather", or sqlite (default is "parquet").
cache_path	(Optional) Path to where to cache the table permanently. By default, the data is cached in the path specified by 'Sys.getenv("CANSIM_CACHE_PATH")', if this is set. Otherwise it will use 'tempdir()'.

Examples

```
## Not run:
cansim_repartition_cached_table("34-10-0013", new_partitioning=c("GeoUID"))

## End(Not run)
```

categories_for_level *Use metadata to extract categories for column of specific level*

Description

For tables with data with hierarchical categories, metadata containing hierarchy level descriptions is used to extract categories at a specified level of hierarchy only.

Usage

```
categories_for_level(
  data,
  column_name,
  level = NA,
  strict = FALSE,
  remove_duplicates = TRUE
)
```

Arguments

data	data table object as returned from get_cansim()
column_name	the quoted name of the column to extract categories from
level	the hierarchy level depth to which to extract categories, where 0 is top category
strict	(default FALSE) when TRUE will only extract that specific hierarchy level
remove_duplicates	(default TRUE) When set to TRUE higher level grouping categories already captured by lower level hierarchy data will be removed

Value

A vector of categories

Examples

```
## Not run:
data <- get_cansim("16-10-0117")
categories_for_level(data,"North American Industry Classification System (NAICS)",level=2)

## End(Not run)
```

`collect_and_normalize` *Collect data from a parquet, feather or sqlite query and normalize cansim table output*

Description

Collect data from a parquet, feather or sqlite query and normalize cansim table output

Usage

```
collect_and_normalize(
  connection,
  replacement_value = "val_norm",
  normalize_percent = TRUE,
  default_month = "07",
  default_day = "01",
  factors = TRUE,
  strip_classification_code = FALSE,
  disconnect = FALSE
)
```

Arguments

<code>connection</code>	A connection to a local arrow connection as returned by <code>get_cansim_connection</code> , possibly with filters or other dplyr verbs applied
<code>replacement_value</code>	(Optional) the name of the column the manipulated value should be returned in. Defaults to adding the 'val_norm' value field.
<code>normalize_percent</code>	(Optional) When true (the default) normalizes percentages by changing them to rates
<code>default_month</code>	The default month that should be used when creating Date objects for annual data (default set to "07")
<code>default_day</code>	The default day of the month that should be used when creating Date objects for monthly data (default set to "01")
<code>factors</code>	(Optional) Logical value indicating if dimensions should be converted to factors. (Default set to FALSE).
<code>strip_classification_code</code>	(Optional) Logical value indicating if classification code should be stripped from names. (Default set to false).
<code>disconnect</code>	(Optional) Only used when format is sqlite. Logical value to indicate if the SQLite database connection should be disconnected. (Default is FALSE)

Value

A tibble with the collected and normalized data

Examples

```
## Not run:
library(dplyr)

con <- get_cansim_connection("34-10-0013")
data <- con %>%
  filter(GEO=="Ontario") %>%
  collect_and_normalize()

## End(Not run)
```

disconnect_cansim_connection

Disconnect from a cansim connection

Description

Closes the database connection behind a table retrieved with `get_cansim_connection(..., format="sqlite")`. Parquet and feather connections hold no connection to close and are left alone, so code that does not know which format it was handed can close it either way.

Usage

```
disconnect_cansim_connection(connection)
```

Arguments

`connection` A connection to a cansim table as returned by `get_cansim_connection`

Value

‘NULL’

Examples

```
## Not run:
con <- get_cansim_connection("34-10-0013", format="sqlite")
disconnect_cansim_connection(con)

## End(Not run)
```

get_cansim

*Retrieve a Statistics Canada data table using NDM catalogue number***Description**

Retrieves a data table using an NDM catalogue number as a tidy data frame. Retrieved table data is cached for the duration of the current R session only by default.

Usage

```
get_cansim(
  cansimTableNumber,
  language = "english",
  refresh = FALSE,
  timeout = 200,
  factors = TRUE,
  default_month = "07",
  default_day = "01"
)
```

Arguments

cansimTableNumber	the NDM table number to load
language	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored
refresh	(Optional) When set to TRUE, forces a reload of data table (default is FALSE)
timeout	(Optional) Number of seconds StatCan is allowed to go without sending data before the download is abandoned, to work around scenarios where StatCan servers drop the network connection. This does not limit how long a download may take overall, a transfer that keeps delivering data is left alone. StatCan prepares a whole response before sending any of it, which for large requests can take the better part of a minute, so values much below the default of 200 risk cutting off legitimate requests.
factors	(Optional) Logical value indicating if dimensions should be converted to factors. (Default set to TRUE).
default_month	The default month that should be used when creating Date objects for annual data (default set to "07")
default_day	The default day of the month that should be used when creating Date objects for monthly data (default set to "01") Set to higher values for large tables and slow network connection. (Default is 200).

Value

A tibble with StatCan Table data and added Date column with inferred date objects and added val_norm column with normalized value from the VALUE column.

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
## Not run:
get_cansim("34-10-0013")

## End(Not run)
```

```
get_cansim_changed_series_data_for_coordinates
```

Retrieve data for series that changed, by table and coordinate

Description

Retrieve the data points Statistics Canada changed for the given coordinates of a table. Coordinates among the ones asked about that did not change contribute no rows, and if none of them changed the result is an empty table rather than an error. The StatCan API can only process 300 coordinates at a time, if more than 300 coordinates are specified the function will batch the requests to the API.

Usage

```
get_cansim_changed_series_data_for_coordinates(
  cansimTableNumber,
  coordinates,
  language = "english",
  timeout = 200,
  factors = TRUE,
  default_month = "07",
  default_day = "01"
)
```

Arguments

cansimTableNumber	The table number the coordinates belong to
coordinates	The coordinates to retrieve changed data for
language	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored

timeout	(Optional) Number of seconds StatCan is allowed to go without sending data before the download is abandoned, to work around scenarios where StatCan servers drop the network connection. This does not limit how long a download may take overall, a transfer that keeps delivering data is left alone. StatCan prepares a whole response before sending any of it, which for large requests can take the better part of a minute, so values much below the default of 200 risk cutting off legitimate requests.
factors	(Optional) Logical value indicating if dimensions should be converted to factors. (Default set to TRUE).
default_month	The default month that should be used when creating Date objects for annual data (default set to "07")
default_day	The default day of the month that should be used when creating Date objects for monthly data (default set to "01")

Value

A tibble with the changed data for the specified coordinates

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
## Not run:
get_cansim_changed_series_data_for_coordinates("34-10-0013", "1.1")

## End(Not run)
```

```
get_cansim_changed_series_data_for_vectors
```

Retrieve data for series that changed, by vector

Description

Retrieve the data points Statistics Canada changed for the given vectors. Series among the ones asked about that did not change contribute no rows, and if none of them changed the result is an empty table rather than an error. The StatCan API can only process 300 vectors at a time, if more than 300 vectors are specified the function will batch the requests to the API.

Usage

```
get_cansim_changed_series_data_for_vectors(
  vectors,
  language = "english",
  timeout = 200,
  factors = TRUE,
  default_month = "07",
  default_day = "01"
)
```

Arguments

vectors	The list of vectors to retrieve changed data for
language	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored
timeout	(Optional) Number of seconds StatCan is allowed to go without sending data before the download is abandoned, to work around scenarios where StatCan servers drop the network connection. This does not limit how long a download may take overall, a transfer that keeps delivering data is left alone. StatCan prepares a whole response before sending any of it, which for large requests can take the better part of a minute, so values much below the default of 200 risk cutting off legitimate requests.
factors	(Optional) Logical value indicating if dimensions should be converted to factors. (Default set to TRUE).
default_month	The default month that should be used when creating Date objects for annual data (default set to "07")
default_day	The default day of the month that should be used when creating Date objects for monthly data (default set to "01")

Value

A tibble with the changed data for the specified vector(s)

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
## Not run:
get_cansim_changed_series_data_for_vectors("v41690973")

## End(Not run)
```

```
get_cansim_changed_series_list
```

Retrieve the series that changed today

Description

Retrieve the list of data series Statistics Canada changed today, as vectors together with the table and coordinate they belong to. Where `get_cansim_changed_tables()` reports which tables were touched, this reports the individual series inside them, which is the finer grained way to decide what needs re-downloading.

Usage

```
get_cansim_changed_series_list(timeout = 600)
```

Arguments

`timeout` (Optional) Number of seconds StatCan is allowed to go without sending data before the download is abandoned. The default is set high because this method is silent while it works.

Details

StatCan serves this for the current day only and fills it during the daily update window that ends at 8:30am Eastern. Unlike the `changed_tables` method there is no way to ask for an earlier day, StatCan answers a request naming a date with an HTTP 404.

How long this takes depends entirely on how much StatCan released that morning. The method takes no parameters, so there is no way to ask for a smaller slice of a busy day, and StatCan works out a whole response before sending any of it. On a quiet day the answer arrives in well under a second; on a heavy one the series changing can number in the hundreds of thousands and the request has been seen to outlive StatCan's own gateway, coming back as an HTTP 504 after some nine minutes of silence. That is a limit at StatCan's end which raising `timeout` cannot lift, so on such a day `get_cansim_changed_tables()` is the question worth asking instead.

Value

A tibble with one row per changed series, carrying the vector, the table number, the coordinate and the release time

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
## Not run:
get_cansim_changed_series_list()

## End(Not run)
```

```
get_cansim_changed_tables
```

Retrieve a list of modified tables since a given date

Description

Retrieve a list of tables that have been modified or updated since the specified date.

Usage

```
get_cansim_changed_tables(start_date, end_date = NULL)
```

Arguments

start_date	Starting date in YYYY-MM-DD format to look for changes that changed on or after that date
end_date	Optional end date in YYYY-MM-DD format to look for changes that changed on or before that date, default is same as start date

Value

A tibble with Statistics Canada data table product ids and their release times

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_changed_tables("2018-08-01")
```

```
get_cansim_code_set    Get NDM code sets
```

Description

Useful to get a list of surveys or subjects and used internally

Usage

```
get_cansim_code_set(
  code_set = c("scalar", "frequency", "symbol", "status", "uom", "survey", "subject",
    "wdsResponseStatus"),
  refresh = FALSE
)
```

Arguments

code_set	the code set to retrieve.
refresh	Default is FALSE, repeated calls during the same session will hit the cached data. To refresh the code list during a running R session set to TRUE

Value

A tibble with english and french labels for the given code set

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_code_set("survey")
```

get_cansim_column_categories

Retrieve Statistics Canada data table categories for a specific column

Description

Returns table column details given an NDM table number in English or French. Retrieved table information data is cached for the duration of the R session only.

Usage

```
get_cansim_column_categories(
  cansimTableNumber,
  column,
  language = "english",
  refresh = FALSE,
  timeout = 200
)
```

Arguments

cansimTableNumber	the NDM table number to load
column	the specified column for which to retrieve category information for
language	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored
refresh	(Optional) When set to TRUE, forces a reload of data table (default is FALSE)
timeout	(Optional) Number of seconds StatCan is allowed to go without sending data before the download is abandoned, to work around scenarios where StatCan servers drop the network connection. This does not limit how long a download may take overall, a transfer that keeps delivering data is left alone. StatCan prepares a whole response before sending any of it, which for large requests can take the better part of a minute, so values much below the default of 200 risk cutting off legitimate requests.

Value

A tibble with detailed information on StatCan table categories for the specified field

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_column_categories("34-10-0013", "Geography")
```

```
get_cansim_column_list
```

Retrieve Statistics Canada data table column list

Description

Returns table column details given an NDM table number in English or French. Retrieved table information data is cached for the duration of the R session only.

Usage

```
get_cansim_column_list(
  cansimTableNumber,
  language = "english",
  refresh = FALSE,
  timeout = 200
)
```

Arguments

cansimTableNumber	the NDM table number to load
language	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored
refresh	(Optional) When set to TRUE, forces a reload of data table (default is FALSE)
timeout	(Optional) Number of seconds StatCan is allowed to go without sending data before the download is abandoned, to work around scenarios where StatCan servers drop the network connection. This does not limit how long a download may take overall, a transfer that keeps delivering data is left alone. StatCan prepares a whole response before sending any of it, which for large requests can take the better part of a minute, so values much below the default of 200 risk cutting off legitimate requests.

Value

A tibble listing the column names of the StatCan table.

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_column_list("34-10-0013")
```

`get_cansim_connection` *Retrieve a Statistics Canada data table using NDM catalogue number as parquet, feather, or sqlite database connection*

Description

Retrieves a data table using an NDM catalogue number as parquet, feather, or SQLite database connection. Retrieved table data is cached permanently if a cache path is supplied or for duration of the current R session. If the table is cached the function will check if a newer version is available and emit a warning message if the cached table is out of date.

Usage

```
get_cansim_connection(
  cansimTableNumber,
  language = "english",
  format = "parquet",
  partitioning = c(),
  refresh = FALSE,
  timeout = 1000,
  cache_path = Sys.getenv("CANSIM_CACHE_PATH")
)
```

Arguments

<code>cansimTableNumber</code>	the NDM table number to load
<code>language</code>	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored
<code>format</code>	(Optional) The format of the data table to retrieve. Either "parquet", "feather", or sqlite (default is "parquet").
<code>partitioning</code>	(Optional) Partition columns to use for parquet or feather formats.
<code>refresh</code>	(Optional) Valid options are FALSE (the default), TRUE, and "auto". When set to TRUE, forces a reload of data table, when set to "auto" it will refresh the table by downloading the newest version from StatCan if the table is out of date. If set to FALSE and the table is out of date a warning will be emitted to alert the user that the data is outdated.
<code>timeout</code>	(Optional) Number of seconds StatCan is allowed to go without sending data before the download is abandoned, to work around scenarios where StatCan servers drop the network connection. This does not limit how long a download may take overall, a transfer that keeps delivering data is left alone. StatCan prepares a whole response before sending any of it, which for large requests can take the better part of a minute, so values much below the default of 200 risk cutting off legitimate requests.

cache_path (Optional) Path to where to cache the table permanently. By default, the data is cached in the path specified by 'Sys.getenv('CANSIM_CACHE_PATH')', if this is set. Otherwise it will use 'tempdir()'.

Value

A database connection to a local parquet, feather, or sqlite database with the StatCan Table data. The data frames after calling 'collect()' or 'collect_and_normalize()' are identical up to possibly different row order.

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
## Not run:
con <- get_cansim_connection("34-10-0013")

# Work with the data connection
glimpse(con)

## End(Not run)
```

get_cansim_cube_metadata

Retrieve table metadata from Statistics Canada API

Description

Retrieves table metadata given an input table number or vector of table numbers using either the new or old table number format. Patience is suggested as the Statistics Canada API can be very slow. The 'list_cansim_tables()' function can be used as an alternative to retrieve a (cached) list of CANSIM tables with (more limited) metadata.

Usage

```
get_cansim_cube_metadata(cansimTableNumber, type = "overview", refresh = FALSE)
```

Arguments

cansimTableNumber	A new or old CANSIM/NDM table number or a vector of table numbers
type	Which type of metadata to get, options are "overview", "members", "notes", or "corrections".
refresh	Refresh the data from the Statistics Canada API

Value

a tibble containing the table metadata. When several table numbers are given, the metadata for all tables is retrieved in a single API call and the results are stacked. Types other than "overview" carry no table identifier of their own, for those a 'cansimTableNumber' column is added to identify the table.

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_cube_metadata("34-10-0013")
```

```
get_cansim_data_for_table_coord_periods
```

Retrieve data for specified Statistics Canada data product for last N periods for specific coordinates

Description

Allows for the retrieval of data for a Statistics Canada data table with specific table and coordinates. This allows partial targeted download of tables and can be effectively combined with the `get_cansim_table_template` function to help pinpoint data series of interest. The StatCan API can only process 300 coordinates at a time, if more than 300 coordinates are specified the function will batch the requests to the API.

Usage

```
get_cansim_data_for_table_coord_periods(
  tableCoordinates,
  periods = NULL,
  language = "english",
  refresh = FALSE,
  timeout = 200,
  factors = TRUE,
  default_month = "07",
  default_day = "01"
)
```

Arguments

`tableCoordinates`

Either a list with vectors of coordinates by table number, or a (filtered) data frame as returned by `get_cansim_table_template`.

`periods`

Optional numeric value for number of latest periods to retrieve data for, default is NULL in which case data for all periods is downloaded. Alternatively this can be specified by coordinate if `tableCoordinates` is a data frame, this argument will be ignored if that data frame as a "periods" column.

language	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored
refresh	(Optional) When set to TRUE, forces a reload of data table (default is FALSE)
timeout	(Optional) Number of seconds StatCan is allowed to go without sending data before the download is abandoned, to work around scenarios where StatCan servers drop the network connection. This does not limit how long a download may take overall, a transfer that keeps delivering data is left alone. StatCan prepares a whole response before sending any of it, which for large requests can take the better part of a minute, so values much below the default of 200 risk cutting off legitimate requests.
factors	(Optional) Logical value indicating if dimensions should be converted to factors. (Default set to TRUE).
default_month	The default month that should be used when creating Date objects for annual data (default set to "07")
default_day	The default day of the month that should be used when creating Date objects for monthly data (default set to "01")

Value

A tibble with data matching specified coordinate and period input arguments

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_data_for_table_coord_periods(list("35-10-0003"=c("1.1", "1.12")), periods=3)
```

```
get_cansim_key_release_schedule
```

Major economic indicator release schedule

Description

Returns every release date of major economic indicators since March 14, 2012. This also includes scheduled future releases.

Usage

```
get_cansim_key_release_schedule()
```

Value

a tibble with data, and details for major economic indicator release

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_key_release_schedule()
```

```
get_cansim_series_info_cube_coord
```

Retrieve series info for given table id and coordinates

Description

Retrieves series information by coordinates

Usage

```
get_cansim_series_info_cube_coord(
  cansimTableNumber,
  coordinates,
  timeout = 1000,
  refresh = FALSE
)
```

Arguments

cansimTableNumber	A new or old CANSIM/NDM table number, coordinates are specific to a single table
coordinates	A vector of coordinates
timeout	(Optional) Number of seconds StatCan is allowed to go without sending data before the call is abandoned. This does not limit how long the call may take overall, a response that keeps arriving is left alone.
refresh	Refresh the data from the Statistics Canada API

Value

a tibble containing the series information for the given coordinates

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_series_info_cube_coord("34-10-0013", c("1.1.1.1.1.1", "2.1.1.1.1.1"))
```

get_cansim_table_info *Retrieve Statistics Canada data table information*

Description

Returns table information given an NDM table catalogue number in English or French. Retrieved table information data is cached for the duration of the R session only.

Usage

```
get_cansim_table_info(
  cansimTableNumber,
  language = "english",
  refresh = FALSE,
  timeout = 200
)
```

Arguments

cansimTableNumber	the NDM table number to load
language	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored
refresh	(Optional) When set to TRUE, forces a reload of data table (default is FALSE)
timeout	(Optional) Number of seconds StatCan is allowed to go without sending data before the download is abandoned, to work around scenarios where StatCan servers drop the network connection. This does not limit how long a download may take overall, a transfer that keeps delivering data is left alone. StatCan prepares a whole response before sending any of it, which for large requests can take the better part of a minute, so values much below the default of 200 risk cutting off legitimate requests.

Value

A tibble with the table overview information

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_table_info("34-10-0013")
```

`get_cansim_table_last_release_date`*Get the latest release data for a StatCan table, if available*

Description

This can be used to check when a table has last been updated.

Usage

```
get_cansim_table_last_release_date(cansimTableNumber)
```

Arguments

cansimTableNumber
the NDM table number

Value

A datetime object if a release data is available, NULL otherwise.

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_table_last_release_date("34-10-0013")
```

`get_cansim_table_notes`*Retrieve Statistics Canada data table notes and column categories*

Description

Returns table notes given an NDM table number in English or French. Retrieved table information data is cached for the duration of the R session only.

Usage

```
get_cansim_table_notes(  
  cansimTableNumber,  
  language = "english",  
  refresh = FALSE,  
  timeout = 200  
)
```

Arguments

cansimTableNumber	the NDM table number to load
language	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored
refresh	(Optional) When set to TRUE, forces a reload of data table (default is FALSE)
timeout	(Optional) Number of seconds StatCan is allowed to go without sending data before the download is abandoned, to work around scenarios where StatCan servers drop the network connection. This does not limit how long a download may take overall, a transfer that keeps delivering data is left alone. StatCan prepares a whole response before sending any of it, which for large requests can take the better part of a minute, so values much below the default of 200 risk cutting off legitimate requests.

Value

A tibble with table notes.

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_table_notes("34-10-0013")
```

```
get_cansim_table_overview
```

Retrieve Statistics Canada data table overview text

Description

Prints table overview information as console output. In order to display table overview information, the selected CANSIM table must be loaded entirely to display overview information. Overview information is printed in console an in English or French, as specified.

Usage

```
get_cansim_table_overview(
  cansimTableNumber,
  language = "english",
  refresh = FALSE
)
```

Arguments

cansimTableNumber	the NDM table number to load
language	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored
refresh	(Optional) When set to TRUE, forces a reload of data table (default is FALSE)

Value

none

Nothing is printed if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_table_overview("34-10-0013")
```

```
get_cansim_table_short_notes
```

Retrieve Statistics Canada data table short notes

Description

Returns table notes given an NDM table number in English or French. Retrieved table information data is cached for the duration of the R session only.

Usage

```
get_cansim_table_short_notes(
  cansimTableNumber,
  language = "english",
  refresh = FALSE,
  timeout = 200
)
```

Arguments

cansimTableNumber	the NDM table number to load
language	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored
refresh	(Optional) When set to TRUE, forces a reload of data table (default is FALSE)

timeout (Optional) Number of seconds StatCan is allowed to go without sending data before the download is abandoned, to work around scenarios where StatCan servers drop the network connection. This does not limit how long a download may take overall, a transfer that keeps delivering data is left alone. StatCan prepares a whole response before sending any of it, which for large requests can take the better part of a minute, so values much below the default of 200 risk cutting off legitimate requests.

Value

A tibble with the StatCan Notes for the table

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_table_short_notes("34-10-0013")
```

```
get_cansim_table_subject
```

Retrieve Statistics Canada data table subject detail

Description

Returns table subject detail given an NDM table number in English or French. Retrieved table information data is cached for the duration of the R session only.

Usage

```
get_cansim_table_subject(
  cansimTableNumber,
  language = "english",
  refresh = FALSE,
  timeout = 200
)
```

Arguments

cansimTableNumber	the NDM table number to load
language	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored
refresh	(Optional) When set to TRUE, forces a reload of data table (default is FALSE)

timeout (Optional) Number of seconds StatCan is allowed to go without sending data before the download is abandoned, to work around scenarios where StatCan servers drop the network connection. This does not limit how long a download may take overall, a transfer that keeps delivering data is left alone. StatCan prepares a whole response before sending any of it, which for large requests can take the better part of a minute, so values much below the default of 200 risk cutting off legitimate requests.

Value

A tibble with the table subject code and name.

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_table_subject("34-10-0013")
```

```
get_cansim_table_survey
```

Retrieve Statistics Canada data table survey detail

Description

Returns table survey detail given an NDM table number in English or French. Retrieved table information data is cached for the duration of the R session only.

Usage

```
get_cansim_table_survey(
  cansimTableNumber,
  language = "english",
  refresh = FALSE,
  timeout = 200
)
```

Arguments

cansimTableNumber	the NDM table number to load
language	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored
refresh	(Optional) When set to TRUE, forces a reload of data table (default is FALSE)

timeout (Optional) Number of seconds StatCan is allowed to go without sending data before the download is abandoned, to work around scenarios where StatCan servers drop the network connection. This does not limit how long a download may take overall, a transfer that keeps delivering data is left alone. StatCan prepares a whole response before sending any of it, which for large requests can take the better part of a minute, so values much below the default of 200 risk cutting off legitimate requests.

Value

A tibble with the table survey code and name

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_table_survey("34-10-0013")
```

```
get_cansim_table_template
```

Retrieve table template from Statistics Canada API

Description

A table template consists of the dimensions and members and coordinates of a table that can be used to explore and filter table data before downloading subsets of the table. To add vector Ids to (a possibly filtered) template the 'add_cansim_vectors_to_template' function can be used.

Usage

```
get_cansim_table_template(
  cansimTableNumber,
  language = "english",
  refresh = FALSE
)
```

Arguments

cansimTableNumber	A new or old CANSIM/NDM table number or a vector of table numbers
language	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored
refresh	Refresh the data from the Statistics Canada API

Value

a tibble containing the table template, with a 'cansimTableName' column identifying the table. When several table numbers are given, the templates are stacked and columns for dimensions that only appear in some of the tables are filled with 'NA' for the other tables.

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_table_template("34-10-0013")
```

get_cansim_table_url	<i>Retrieve a Statistics Canada data table URL given a table number</i>
----------------------	---

Description

Retrieve URL of a table from the API given a table number. Offers a more stable approach than manually guessing the URL of the table.

Usage

```
get_cansim_table_url(cansimTableName, language = "english")
```

Arguments

cansimTableName	the NDM table number to load
language	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored

Value

String object containing URL for specified table number

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_table_url("34-10-0013")  
get_cansim_table_url("34-10-0013", language = "fr")
```

get_cansim_vector	<i>Retrieve data for a Statistics Canada data vector released within a given time frame</i>
-------------------	---

Description

Allows for the retrieval of data for specified vector series for a given time window. Accessing data by vector allows for targeted extraction of time series. Discovering vectors of interest can be achieved using the StatCan table web interface or using `get_cansim_table_template` function to help pinpoint data series of interest, and then chaining the `add_cansim_vectors_to_template` function to add cansim vector information to the template data. The StatCan API can only process 300 coordinates at a time, if more than 300 coordinates are specified the function will batch the requests to the API.

Usage

```
get_cansim_vector(
  vectors,
  start_time = as.Date("1800-01-01"),
  end_time = Sys.time(),
  use_ref_date = TRUE,
  language = "english",
  refresh = FALSE,
  timeout = 200,
  factors = TRUE,
  default_month = "07",
  default_day = "01"
)
```

Arguments

vectors	The list of vectors to retrieve
start_time	Starting date in YYYY-MM-DD format, applies to REF_DATE or releaseTime, depending on use_ref_date parameter
end_time	Set an optional end time filter in YYYY-MM-DD format (defaults to current system time)
use_ref_date	Optional, TRUE by default. When set to TRUE, uses REF_DATE of vector data to filter, otherwise it uses StatisticsCanada's releaseDate value for filtering the specified vectors.
language	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored
refresh	(Optional) When set to TRUE, forces a reload of data table (default is FALSE)

timeout	(Optional) Number of seconds StatCan is allowed to go without sending data before the download is abandoned, to work around scenarios where StatCan servers drop the network connection. This does not limit how long a download may take overall, a transfer that keeps delivering data is left alone. StatCan prepares a whole response before sending any of it, which for large requests can take the better part of a minute, so values much below the default of 200 risk cutting off legitimate requests.
factors	(Optional) Logical value indicating if dimensions should be converted to factors. (Default set to TRUE).
default_month	The default month that should be used when creating Date objects for annual data (default set to "07")
default_day	The default day of the month that should be used when creating Date objects for monthly data (default set to "01")

Value

A tibble with data for vectors released between start and end time

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_vector("v41690973", "2015-01-01")
```

```
get_cansim_vector_for_latest_periods
```

Retrieve data for specified Statistics Canada data vector(s) for last N periods

Description

Allows for the retrieval of data for specified vector series for the N most-recently released periods. Accessing data by vector allows for targeted extraction of time series. Discovering vectors of interest can be achieved using the StatCan table web interface or using `get_cansim_table_template` function to help pinpoint data series of interest, and then chaining the `add_cansim_vectors_to_template` function to add cansim vector information to the template data. The StatCan API can only process 300 coordinates at a time, if more than 300 coordinates are specified the function will batch the requests to the API.

Usage

```
get_cansim_vector_for_latest_periods(
  vectors,
  periods = NULL,
  language = "english",
  refresh = FALSE,
```

```

    timeout = 200,
    factors = TRUE,
    default_month = "07",
    default_day = "01"
  )

```

Arguments

vectors	The list of vectors to retrieve
periods	Numeric value for number of latest periods to retrieve data for, but default all data is retrieved.
language	"english" (the default) or "french". Short forms such as "en", "eng", "fr" or "fra" are accepted, as are the French names "anglais" and "français"; case and accents are ignored
refresh	(Optional) When set to TRUE, forces a reload of data table (default is FALSE)
timeout	(Optional) Number of seconds StatCan is allowed to go without sending data before the download is abandoned, to work around scenarios where StatCan servers drop the network connection. This does not limit how long a download may take overall, a transfer that keeps delivering data is left alone. StatCan prepares a whole response before sending any of it, which for large requests can take the better part of a minute, so values much below the default of 200 risk cutting off legitimate requests.
factors	(Optional) Logical value indicating if dimensions should be converted to factors. (Default set to TRUE).
default_month	The default month that should be used when creating Date objects for annual data (default set to "07")
default_day	The default day of the month that should be used when creating Date objects for monthly data (default set to "01")

Value

A tibble with data for specified vector(s) for the last N periods

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_vector_for_latest_periods("v41690973", 10)
```

`get_cansim_vector_info`*Retrieve metadata for specified Statistics Canada data vectors*

Description

Allows for the retrieval of metadata for Statistics Canada data vectors

Usage

```
get_cansim_vector_info(vectors)
```

Arguments

`vectors` a vector of cansim vectors

Value

A tibble with metadata for selected vectors

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
get_cansim_vector_info("v41690973")
```

`list_cansim_cached_tables`*List cached cansim arrow and SQLite databases*

Description

List cached cansim arrow and SQLite databases

Usage

```
list_cansim_cached_tables(  
  cache_path = Sys.getenv("CANSIM_CACHE_PATH"),  
  refresh = FALSE  
)
```

Arguments

`cache_path` Optional, default value is 'Sys.getenv('CANSIM_CACHE_PATH')'.
`refresh` Optional, refresh the last updated date of cached cansim tables

Value

A tibble with the list of all tables that are currently cached at the given cache path.

Examples

```
## Not run:
list_cansim_cached_tables()

## End(Not run)
```

list_cansim_cubes	<i>Get overview list for all Statistics Canada data cubes</i>
-------------------	---

Description

Generates an overview table containing metadata of available Statistics Canada data cubes.

Usage

```
list_cansim_cubes(lite = FALSE, refresh = FALSE, quiet = FALSE)
```

Arguments

lite	Get the version without cube dimensions and comments for faster retrieval, default is FALSE.
refresh	Default is FALSE, repeated calls during the same session will hit the cached data.
quiet	Optional, suppress messages To refresh the code list during a running R session set to TRUE

Value

A tibble with available Statistics Canada data cubes, including NDM table number, cube title, start and end dates, achieve status, subject and survey codes, frequency codes and a list of cube dimensions.

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
## Not run:
list_cansim_cubes()

## End(Not run)
```

remove_cansim_cached_tables

Remove cached cansim SQLite and parquet database

Description

Remove cached cansim SQLite and parquet database

Usage

```
remove_cansim_cached_tables(
  cansimTableName,
  format = c("parquet", "feather", "sqlite"),
  language = NULL,
  cache_path = Sys.getenv("CANSIM_CACHE_PATH")
)
```

Arguments

cansimTableName	Vector of the table(s) to be removed, or a (filtered) table as returned by ‘list_cansim_cached_tables’ with the list of tables to be removed.
format	Format of cache to remove, possible values are “parquet”, “feather” or “sqlite” or a subset of these (the default is all of these)
language	Language for which to remove the cached data, named as in get_cansim(). If unspecified (‘NULL’) tables for all languages will be removed.
cache_path	Optional, default value is ‘Sys.getenv(‘CANSIM_CACHE_PATH’)’

Value

‘NULL’

Examples

```
## Not run:
con <- get_cansim_connection("34-10-0013", format="parquet")
remove_cansim_cached_tables("34-10-0013", format="parquet")

## End(Not run)
```

search_cansim_cubes	<i>Search through Statistics Canada data cubes</i>
---------------------	--

Description

Searches through Statistics Canada data cubes using a search term.

Usage

```
search_cansim_cubes(search_term, refresh = FALSE)
```

Arguments

search_term	User-supplied search term used to find Statistics Canada data cubes with matching titles, table numbers, subject and survey codes.
refresh	Default is FALSE. The underlying cube list is cached for the duration of the R sessions and will regenerate the cube list if set to TRUE

Value

A tibble with available Statistics Canada data cubes, listing title, Statistics Canada data cube catalogue number, deprecated CANSIM table number, survey and subject.

Returns NULL if the data could not be retrieved because StatCan is unavailable.

Examples

```
## Not run:
search_cansim_cubes("Labour force")

## End(Not run)
```

set_cansim_cache_path	<i>Set persistent cansim cache location</i>
-----------------------	---

Description

Cansim provides session caching for retrieved data. This function will create a persistent cache across sessions for data accessed via 'get_cansim_connection' and caches data in a database across sessions..

Usage

```
set_cansim_cache_path(cache_path, overwrite = FALSE, install = FALSE)
```

Arguments

cache_path	a local directory to use for saving cached data
overwrite	Option to overwrite any existing cache path already stored locally.
install	Option to install permanently for use across sessions.

Examples

```
## Not run:
set_cansim_cache_path("~/cansim_cache")

# This will set the cache path permanently until overwritten again
set_cansim_cache_path("~/cancensus_cache", install = TRUE)

## End(Not run)
```

```
show_cansim_cache_path
```

View saved cache directory path

Description

View saved cache path'

Usage

```
show_cansim_cache_path()
```

Examples

```
show_cansim_cache_path()
```

```
view_cansim_webpage
```

View CANSIM table or vector information in browser

Description

Opens CANSIM table or vector on Statistics Canada's website using default browser. This may be useful for getting further info on CANSIM table and survey methods.

Usage

```
view_cansim_webpage(cansimTableNumber = NULL)
```

Arguments

`cansimTableNumber`

CANSIM or NDM table number or cansim vectors with "v" prefix. If no number is provided, the vector search page on the Statistic Canada website will be opened.

Value

none

Examples

```
## Not run:  
view_cansim_webpage("34-10-0013")  
  
## End(Not run)
```

Index

[add_cansim_vectors_to_template](#), 3
[add_provincial_abbreviations](#), 4

[cansim_old_to_new](#), 4
[cansim_repartition_cached_table](#), 5
[categories_for_level](#), 6
[collect_and_normalize](#), 7

[disconnect_cansim_connection](#), 8

[get_cansim](#), 9
[get_cansim_changed_series_data_for_coordinates](#),
10
[get_cansim_changed_series_data_for_vectors](#),
11
[get_cansim_changed_series_list](#), 12
[get_cansim_changed_tables](#), 13
[get_cansim_code_set](#), 14
[get_cansim_column_categories](#), 15
[get_cansim_column_list](#), 16
[get_cansim_connection](#), 17
[get_cansim_cube_metadata](#), 18
[get_cansim_data_for_table_coord_periods](#),
19
[get_cansim_key_release_schedule](#), 20
[get_cansim_series_info_cube_coord](#), 21
[get_cansim_table_info](#), 22
[get_cansim_table_last_release_date](#), 23
[get_cansim_table_notes](#), 23
[get_cansim_table_overview](#), 24
[get_cansim_table_short_notes](#), 25
[get_cansim_table_subject](#), 26
[get_cansim_table_survey](#), 27
[get_cansim_table_template](#), 28
[get_cansim_table_url](#), 29
[get_cansim_vector](#), 30
[get_cansim_vector_for_latest_periods](#),
31
[get_cansim_vector_info](#), 33

[list_cansim_cached_tables](#), 33

[list_cansim_cubes](#), 34

[remove_cansim_cached_tables](#), 35

[search_cansim_cubes](#), 36
[set_cansim_cache_path](#), 36
[show_cansim_cache_path](#), 37

[view_cansim_webpage](#), 37