

Package ‘readabs’

February 1, 2019

Type Package

Title Download and Tidy Time Series Data from the Australian Bureau of Statistics

Version 0.3.0

Maintainer Matt Cowgill <mattcowgill@gmail.com>

Description Downloads, imports, and tidies time series data from the Australian Bureau of Statistics <<https://www.abs.gov.au/>>.

Date 2019-01-29

License MIT + file LICENSE

Encoding UTF-8

LazyData true

Depends R (>= 3.1)

Imports readxl (>= 1.2.0), tibble (>= 1.4.99), dplyr, XML, purrr, tidy, rsdmx, readr

URL <https://github.com/mattcowgill/readabs>

BugReports <https://github.com/mattcowgill/readabs/issues>

RoxygenNote 6.1.1

VignetteBuilder knitr

Suggests knitr, rmarkdown, testthat, RCurl, ggplot2

NeedsCompilation no

Author Matt Cowgill [aut, cre],
Zoe Meers [aut],
Jaron Lee [aut]

Repository CRAN

Date/Publication 2019-02-01 20:20:02 UTC

R topics documented:

extract_abs_sheets	2
get_abs	2
read_abs	3
read_abs_data	4
read_abs_local	4
read_abs_metadata	5
read_abs_sdmx	6
tidy_abs	6
tidy_abs_list	7

Index	8
--------------	----------

extract_abs_sheets	<i>Extract data sheets from an ABS timeseries workbook saved locally as an .xls file.</i>
--------------------	---

Description

Note that this function will not tidy the data for you. Use ‘read_abs_local()’ to import and tidy data from local ABS time series spreadsheets or ‘read_abs()’ to download, import and tidy ABS time series.

Usage

```
extract_abs_sheets(filename, table_title = NULL, path = "data/ABS")
```

Arguments

filename	Filename for an ABS time series spreadsheet (as string), such as "6202002.xls".
table_title	String giving the full title of the ABS table, such as "Table 1. Employed persons, Australia"
path	Local directory in which an ABS time series is stored. Default is "data/ABS"

get_abs	<i>Download, extract, and tidy ABS time series spreadsheets (soft deprecated)</i>
---------	---

Description

get_abs() is deprecated. Please use read_abs() instead. It has identical functionality.

Usage

```
get_abs(...)
```

Arguments

... arguments passed to read_abs()

read_abs *Download, extract, and tidy ABS time series spreadsheets*

Description

read_abs() downloads ABS time series spreadsheets, then extracts the data from those spreadsheets, then tidies the data. The result is a single data frame (tibble) containing tidied data.

Usage

```
read_abs(cat_no = NULL, tables = "all", path = "data/ABS",
         metadata = TRUE, show_progressBars = TRUE)
```

Arguments

cat_no ABS catalogue number, as a string, including the extension. For example, "6202.0".

tables numeric. Time series tables in 'cat_no' to download and extract. Default is "all", which will read all time series in 'cat_no'. Specify 'tables' to download and import specific tables(s) - eg. 'tables = 1' or 'tables = c(1, 5)'.

path Local directory in which to save downloaded ABS time series spreadsheets. Default is "data/ABS"; this subdirectory of your working directory will be created if it does not exist.

metadata logical. If 'TRUE' (the default), a tidy data frame including ABS metadata (series name, table name, etc.) is included in the output. If 'FALSE', metadata is dropped.

show_progressBars TRUE by default. If set to FALSE, progress bars will not be shown when ABS spreadsheets are downloading.

Value

A data frame (tibble) containing the tidied data from the ABS time series table(s).

Examples

```
# Download and tidy all time series spreadsheets from the Wage Price Index (6345.0)
wpi <- read_abs("6345.0")
```

read_abs_data	<i>Extracts ABS time series data from local Excel spreadsheets and converts to long format.</i>
---------------	---

Description

'read_abs_data()' is soft deprecated. Please consider using 'read_abs_local()' to import and tidy locally-stored ABS time series spreadsheets, or 'read_abs()' to download, import, and tidy time series spreadsheets from the ABS website.

Usage

```
read_abs_data(path, sheet)
```

Arguments

path	Filepath to Excel spreadsheet.
sheet	Sheet name or number.

Value

Long-format dataframe

read_abs_local	<i>Read and tidy locally-saved ABS time series spreadsheet(s)</i>
----------------	---

Description

If you need to download and tidy time series data from the ABS, use read_abs(). read_abs_local() imports and tidies data from ABS time series spreadsheets that are already saved to your local drive.

Usage

```
read_abs_local(filenamees = NULL, path = "data/ABS", metadata = TRUE)
```

Arguments

filenamees	character vector of at least one filename of a locally-stored ABS time series spreadsheet. For example, "6202001.xls" or c("6202001.xls", "6202005.xls"). If 'filenamees' is blank, 'read_abs_local()' will attempt to read all .xls files in the directory specified with 'path'. Please note that the directory should be specified with 'path'.
path	path to local directory containing ABS time series file(s). Default is "data/ABS". If nothing is specified in 'filenamees', 'read_abs_local()' will attempt to read all .xls files in the directory specified with 'path'.

metadata logical. If 'TRUE' (the default), a tidy data frame including ABS metadata (series name, table name, etc.) is included in the output. If 'FALSE', metadata is dropped.

Details

Unlike read_abs(), the 'table_title' column in the data frame returned by read_abs_local() is blank. If you require 'table_title', please use read_abs() instead.

Examples

```
# Load and tidy two specified files from the "data/ABS" subdirectory
# of your working directory

lfs <- read_abs_local(c("6202001.xls", "6202005.xls"))
```

read_abs_metadata *Extracts ABS series metadata directly from Excel spreadsheets and converts to long-form.*

Description

Extracts ABS series metadata directly from Excel spreadsheets and converts to long-form.

Usage

```
read_abs_metadata(path, sheet)
```

Arguments

path Filepath to Excel spreadsheet.
sheet Sheet name or number.

Value

Long-form dataframe

read_abs_sdmx	<i>Extracts ABS XML-formatted data using the exported XML link at http://www.abs.gov.au/</i>
---------------	---

Description

Access the sdmx URLs at 'http://www.abs.gov.au/ausstats/abs@.nsf/Lookup/1407.0.55.002Main+Features4User+Guide'

Usage

```
read_abs_sdmx(url)
```

Arguments

url	URL weblink.
-----	--------------

Value

data frame

Examples

```
url <- paste0("http://stat.data.abs.gov.au/restsdmx/sdmx.ashx/GetData/LF/",
             "0.2+3+4+11+13+6+15+14+10.3+1+2.1519+1599.10+20+30.M/",
             "all?startTime=2017-12&endTime=2018-11")
lfs <- read_abs_sdmx(url)
lfs
```

tidy_abs	<i>Tidy ABS time series data.</i>
----------	-----------------------------------

Description

Tidy ABS time series data.

Usage

```
tidy_abs(df, metadata = TRUE)
```

Arguments

df	A data frame containing ABS time series data that has been extracted using <code>extract_abs_sheets</code> .
metadata	logical. If 'TRUE' (the default), a tidy data frame including ABS metadata (series name, table name, etc.) is included in the output. If 'FALSE', metadata is dropped.

Value

data frame (tibble) in long format.

Examples

```
# First extract the data from the local spreadsheet
## Not run: wpi <- extract_abs_sheets("634501.xls")

# Then tidy the data extracted from the spreadsheet. Note that
# \code{extract_abs_sheets()} returns a list of data frames, so we need to
# subset the list.

tidy_wpi <- tidy_abs(wpi[[1]])
```

tidy_abs_list

Tidy multiple dataframes of ABS time series data contained in a list.

Description

Tidy multiple dataframes of ABS time series data contained in a list.

Usage

```
tidy_abs_list(list_of_dfs, metadata = TRUE)
```

Arguments

list_of_dfs	A list of dataframes containing extracted ABS time series data.
metadata	logical. If 'TRUE' (the default), a tidy data frame including ABS metadata (series name, table name, etc.) is included in the output. If 'FALSE', metadata is dropped.

Index

`extract_abs_sheets`, [2](#)

`get_abs`, [2](#)

`read_abs`, [3](#)

`read_abs_data`, [4](#)

`read_abs_local`, [4](#)

`read_abs_metadata`, [5](#)

`read_abs_sdmx`, [6](#)

`tidy_abs`, [6](#)

`tidy_abs_list`, [7](#)