Package 'wcde'

September 5, 2026

Title Download Data from the Wittgenstein Centre Human Capital Data Explorer

Version 0.0.8.1

URL https://guyabel.github.io/wcde/

BugReports https://github.com/guyabel/wcde/issues/

Description Download and plot education specific demographic data from the Wittgenstein Centre for Demography and Human Capital Data Explorer <https://dataexplorer.wittgensteincentre.org/>.

License GPL-3

Encoding UTF-8

LazyData true

Imports dplyr, magrittr, tidyr, progress, countrycode, tibble, purrr, stringr, forcats, RCurl

Depends R (>= 3.5)

Suggests spelling, knitr, rmarkdown, tidyverse, lemon

VignetteBuilder knitr

Language en-US

Config/roxygen2/version 8.0.0

NeedsCompilation no

Author Guy J. Abel [aut, cre, ctb] (ORCID: <https://orcid.org/0000-0002-4893-5687>), Samir K.C. [ctb] (ORCID: <https://orcid.org/0000-0002-5213-9181>), Michaela Potancokova [ctb], Claudia Reiter [ctb] (ORCID: <https://orcid.org/0000-0002-1485-3851>), Andrea Tamburini [ctb], Dilek Yildiz [ctb] (ORCID: <https://orcid.org/0000-0001-6192-0634>)

Maintainer Guy J. Abel <g.j.abel@gmail.com>

Repository CRAN

Date/Publication 2026-09-05 15:50:02 UTC

Contents

edu_group_sum 2
every_other 3
find_indicator 4
get_wcde 4
past_epop 7
wic_col4 8
wic_col6 8
wic_col8 9
wic_indicators 9
wic_locations 10
wic_scenarios 11

Index **12**

`edu_group_sum` *Education group sums*

Description

Cleans epop data, downloaded using the `wcde()` function, for summations of population by 4, 6 or 8 education groups.

Usage

```
edu_group_sum(
  d = NULL,
  n = 4,
  strip_totals = TRUE,
  factor_convert = TRUE,
  year_edu_start = 2020
)
```

Arguments

`d`	Data frame downloaded from the
`n`	Number of education groups (from 4, 6 or 8)
`strip_totals`	Remove total sums in epop column. Will not strip education totals if `year` < `year_edu_start` and `n = 8` as past data on population size by 8 education groups is unavailable.
`factor_convert`	Convert columns that are character strings to factors, with levels based on order of appearance.
`year_edu_start`	Year in which education splits are available for given groupings - in some versions past data is not available for some education groupings. Set to 2020 by default.

Details

Strips the epop data set to relevant rows for the n education groups.

Value

A tibble with the data selected.

Examples

```
library(tidyverse)
past_epop %>%
  filter(year == 2020) %>%
  edu_group_sum()
```

`every_other` *Select every other (nth) element from a vector*

Description

Select every other (nth) element from a vector

Usage

```
every_other(x, n = 2, start = 1, fill = NULL)
```

Arguments

`x`	Vector to select (remove) elements from
`n`	Numeric value for the number of elements to skip. Default is 2, i.e. skips every second element
`start`	Numeric value to indicate which element of the vector to commence from.
`fill`	Character string to be used in place of skipped element. By default is `NULL` and hence skipped elements are removed rather than replaced.

Value

Vector with elements removed

Examples

```
every_other(x = letters)
every_other(LETTERS, n = 3, start = 6)
every_other(x = letters, fill = "")
```

`find_indicator` *Find available indicator code names in the Wittgenstein Centre Human Capital Data Explorer*

Description

Find available indicator code names in the Wittgenstein Centre Human Capital Data Explorer

Usage

```
find_indicator(x)
```

Arguments

`x`	Character string on key word or name related to indicator of potential interest.

Value

A subset of the `wic_indicators` data frame with one or more of the `indicator`, `description` or `definition` columns matching the keyword given to `x`. Use the result in the `indicator` column to input to the `get_wcde` function for downloading data.

Examples

```
find_indicator("education")
find_indicator("migr")
find_indicator("fert")
```

`get_wcde` *Download data from the Wittgenstein Centre Human Capital Data Explorer*

Description

Downloads data from the Wittgenstein Centre Human Capital Data Explorer. Requires a working internet connection.

Usage

```
get_wcde(
  indicator = "pop",
  scenario = 2,
  country_code = NULL,
  country_name = NULL,
  pop_age = c("total", "all"),
  pop_sex = c("total", "both", "all"),
```

```
  pop_edu = c("total", "four", "six", "eight"),
  include_scenario_names = FALSE,
  server = c("iiasa", "github", "1&1", "search-available", "iiasa-local"),
  version = c("wcde-v3", "wcde-v2", "wcde-v1")
)
```

Arguments

indicator	One character string based on the indicator column in the wic_indicators data frame, representing the variable to be downloaded.
scenario	Vector of length one or more with numbers corresponding the scenarios. See details for more information. Defaults to 2 for the SSP2 Medium scenario.
country_code	Vector of length one or more of country numeric codes based on ISO 3 digit numeric values.
country_name	Vector of length one or more of country names. The corresponding country code will be guessed using the countrycodes package.
pop_age	Character string for population age groups if indicator is set to pop. Defaults to no age groups total, but can be set to all.
pop_sex	Character string for population sexes if indicatoris set to pop. Defaults to no sex total, but can be set to both or all.
pop_edu	Character string for population educational attainment if indicator is set to pop. Defaults to total, but can be set to four, six or eight.
include_scenario_names	Logical vector of length one to indicate if to include additional columns for scenario names and short names. FALSE by default.
server	Character string for server to download from. Defaults to iiasa, but can use github or 1&1 if IIASA server is down. Can check availability by setting to search-available. Note, 1&1 only hosts batch files (i.e. measures for all countries)
version	Character string for version of projections to obtain. Defaults to wcde-v3, but can use wcde-v2 or wcde-v1. Scenario and indicator availability vary between versions.

Details

If no country_name or country_code is provided data for all countries and regions are downloaded. A full list of available countries and regions can be found in the wic_locations data frame.

indicator must be set to a value in the first column in the table below of available demographic indicators:

indicator	Indicator Description
pop	Population Size (000's)
bpop	Population Size by Broad Age (000's)
epop	Population Size by Education (000's)
prop	Educational Attainment Distribution

bprop	Educational Attainment Distribution by Broad Age
growth	Average Annual Growth Rate
nirate	Average Annual Rate of Natural Increase
sexratio	Sex Ratio
mage	Population Median Age
tdr	Total Dependency Ratio
ydr	Youth Dependency Ratio
odr	Old-age Dependency Ratio
ryl15	Age When Remaining Life Expectancy is Below 15 years
pryl15	Proportion of Population with a Remaining Life Expectancy below 15 Years
mys	Mean Years of Schooling by Age
bmys	Mean Years of Schooling by Broad Age
ggapmys15	Gender Gap in Mean Years Schooling (15+)
ggapmys25	Gender Gap in Mean Years Schooling (25+)
ggapedu15	Gender Gap in Educational Attainment (15+)
ggapedu25	Gender Gap in Educational Attainment (25+)
tfr	Total Fertility Rate
etfr	Total Fertility Rate by Education
asfr	Age-Specific Fertility Rate
easfr	Age-Specific Fertility Rate by Education
cbr	Crude Birth Rate
macb	Mean Age at Childbearing
emacb	Mean Age at Childbearing by Education
e0	Life Expectancy at Birth
cdr	Crude Death Rate
assr	Age-Specific Survival Ratio
eassr	Age-Specific Survival Ratio by Education
net	Net Migration
netedu	Net Migration Flows by Education
emi	Emigration Flows
imm	Immigration Flows

See `wic_indicators` data frame for more details.

`scenario` must be set to one or values in the first column table below of the available future scenarios:

scenario	description	version
1	Rapid Development (SSP1)	V1, V2, V3
2	Medium (SSP2)	V1, V2, V3
3	Stalled Development (SSP3)	V1, V2, V3
4	Inequality (SSP4)	V1, V3
5	Conventional Development (SSP5)	V1, V3
20	Medium - Constant Enrollment Rate (SSP2-CER)	V1
21	Medium - Fast Track Education (SSP2-FT)	V1
22	Medium - Zero Migration (SSP2-ZM)	V2, V3
23	Medium - Double Migration (SSP2-DM)	V2, V3

See `wic_scenarios` data frame for more details.

Value

A tibble with the data selected.

Examples

```
# SSP2 tfr for Austria and Bulgaria
get_wcde(indicator = "tfr", country_code = c(40, 100))

# SSP1 and SSP2 life expectancy for Vietnam and United Kingdom (guessing the country codes)
get_wcde(scenario = c(1, 2), indicator = "e0", country_name = c("Vietnam", "UK"))

# SSP1 and SSP3 population by education for all countries
get_wcde(scenario = c(1, 3), indicator = "tfr")

# population totals (aggregated over age, sex and education)
get_wcde(indicator = "pop", country_name = "Austria")

# population totals by education group
get_wcde(indicator = "pop", country_name = "Austria", pop_edu = "four")

# population totals by age-sex group
get_wcde(indicator = "pop", country_name = "Austria", pop_age = "all", pop_sex = "both")
```

`past_epop` *Past population sizes for all countries by education*

Description

A data set containing population sizes for all countries by education between 1950 and 2020 from Version 3.

Usage

```
past_epop
```

Format

A data frame with 574,968 rows and 7 variables, including:

name Area name

country_code ISO 3 digit country code

year Year of observation from 1950 to 2020 in five-year steps

age Five-year age groups

education Education group

sex Sex

epop Population size in thousands for each age, sex and education group

Source

https://dataexplorer.wittgensteincentre.org/

`wic_col4` *Colours used in Wittgenstein Centre for Demography and Human Capital Data Explorer*

Description

Three sets of colours used for filling education based plots based on the different availability of detailed education categories (four, six or eight groups)

Usage

```
wic_col4
```

Format

A named vector

`wic_col6` *Colours used in Wittgenstein Centre for Demography and Human Capital Data Explorer*

Description

Three sets of colours used for filling education based plots based on the different availability of detailed education categories (four, six or eight groups)

Usage

```
wic_col6
```

Format

A named vector

wic_col8 *Colours used in Wittgenstein Centre for Demography and Human Capital Data Explorer*

Description

Three sets of colours used for filling education based plots based on the different availability of detailed education categories (four, six or eight groups)

Usage

```
wic_col8
```

Format

A named vector

wic_indicators *Indicators used in the Wittgenstein Centre Human Capital Data Explorer*

Description

A data set containing the indicator codes, names and further details used in the Wittgenstein Centre Human Capital Data Explorer

Usage

```
wic_indicators
```

Format

A data frame with 37 rows and 11 variables, including:

indicator Short name of indicator to be used in the `indicator` argument of the `get_wcde()` function

description Brief description of indicator

wcde-v3 Availability in wcde-v3 of `projection-only` or `past-available` (in addition to projections) of indicator. If value is `NA` then indicator not available in version.

wcde-v2 Availability in wcde-v2 of `projection-only` or `past-available` (in addition to projections) of indicator. If value is `NA` then indicator not available in version.

wcde-v1 Availability in wcde-v1 of `projection-only` or `past-available` (in addition to projections) of indicator. If value is `NA` then indicator not available in version.

age Availability of indicator by five-year age groups

bage Availability of indicator by broad age groups

sage Availability of indicator with a new born age group

sex Availability of indicator by sex

edu Availability of indicator by education

period Indicator is a period (flow)

definition_latest Full definition for indicator based on latest available version

Source

https://dataexplorer.wittgensteincentre.org/

`wic_locations` *Locations used in the Wittgenstein Centre Human Capital Data Explorer*

Description

A dataset containing the location codes, names and further details used in the Wittgenstein Centre Human Capital Data Explorer

Usage

```
wic_locations
```

Format

A data frame with 232 rows and 8 variables, including:

name Area name

isono ISO 3 digit country code

continent Continent of country

region UN region of country

dim Category or country/region/area

wcde-v3 Availability of area in Version 3

wcde-v2 Availability of area in Version 2

wcde-v1 Availability of area in Version 1

Source

https://dataexplorer.wittgensteincentre.org/

`wic_scenarios` *Scenarios used in the Wittgenstein Centre Human Capital Data Explorer*

Description

A data set containing the scenario codes, names short names used in the Wittgenstein Centre Human Capital Data Explorer

Usage

```
wic_scenarios
```

Format

A data frame with 9 rows and 6 variables, including:

scenario_name Full scenario name

scenario Code to match help file of `get_wcde` function

scenario_abb Short scenario name

wcde-v3 Availability of area in Version 3

wcde-v2 Availability of area in Version 2

wcde-v1 Availability of area in Version 1

Source

https://dataexplorer.wittgensteincentre.org/

Index

∗ **datasets**
past_epop, 7
wic_col4, 8
wic_col6, 8
wic_col8, 9
wic_indicators, 9
wic_locations, 10
wic_scenarios, 11

edu_group_sum, 2
every_other, 3

find_indicator, 4

get_wcde, 4

past_epop, 7

tibble, *7*

wic_col4, 8
wic_col6, 8
wic_col8, 9
wic_indicators, 9
wic_locations, 10
wic_scenarios, 11