Package 'kayadata'

July 18, 2023

Type Package

Title Kaya Identity Data for Nations and Regions

Version 1.3.0

Date 2023-07-13

Description Provides data for Kaya identity variables (population, gross domestic product, primary energy consumption, and energy-related CO2 emissions) for the world and for individual nations, and utility functions for looking up data, plotting trends of Kaya variables, and plotting the fuel mix for a given country or region. The Kaya identity (Yoichi Kaya and Keiichi Yokobori, ``Environment, Energy, and Economy: Strategies for Sustainability'' (United Nations University Press, 1998) and <https://en.wikipedia.org/wiki/Kaya_identity>) expresses a nation's or region's greenhouse gas emissions in terms of its population, per-capita Gross Domestic Product, the energy intensity of its economy, and the carbon-intensity of its energy supply.

URL https://jonathan-g.github.io/kayadata/,
https://github.com/jonathan-g/kayadata,
https://doi.org/10.5281/zenodo.8144476

BugReports https://github.com/jonathan-g/kayadata/issues

License MIT + file LICENSE

Depends R (>= 3.5), ggplot2 (>= 3.0)

Imports magrittr (>= 1.5), forcats (>= 0.3), dplyr (>= 0.8), tidyr (>= 0.8), stringr (>= 1.3), scales (>= 1.0), purrr (>= 0.3)

Encoding UTF-8

Language en-US

LazyData true

RoxygenNote 7.2.3

Suggests broom (>= 0.5), knitr (>= 1.22), rmarkdown (>= 1.12), testthat (>= 2.0), vdiffr (>= 0.3.1)

VignetteBuilder knitr

NeedsCompilation no

Author Jonathan Gilligan [cre, aut] (<https://orcid.org/0000-0003-1375-6686>)

Maintainer Jonathan Gilligan <jonathan.gilligan@vanderbilt.edu>

Repository CRAN

Date/Publication 2023-07-17 22:40:05 UTC

R topics documented:

`kayadata-package` *kayadata package*

Description

kayadata is a package for working with Kaya identity data for many countries and regions.

The Kaya identity, named for the economist Yoichi Kaya, who introduced it (Kaya, 1998); It decomposes the energy-related carbon dioxide emissions from a nation, region, or the world into the product of four components:

$$F = P \times g \times e \times f,$$

where F is the total emissions, P is the population, g is the per-capita GDP, e is the energy intensity of the economy, and f is the emissions-intensity of the energy supply. (Nakicenovic and Swart, 2000, Ch. 3, p. 105; Raupach *et al*, 2007)

The data in this packages covers 1960-2019 for population and GDP, and 1965-2019 for energy and fossil-fuel CO2 emissions.

The package uses data on population and GDP from the World Bank, using market exchange rates (MER) for GDP because those data go back to 1960. From 1990 onward, Purchasing-Power-Parity (PPP) GDP figures are available as `G_ppp` but using these would require re-calculating `G`, `g`, `e`, and `ef` in the `kaya_data` data frame.

The package uses data on energy consumption and fossil-fuel CO2 emissions from the Energy Institute's 2023 Statistical Review of World Energy

License

The **kayadata** package is open source licensed under the MIT License.

Bug reports

- kayadata issue tracker (https://github.com/jonathan-g/kayadata/issues)

References

Kaya, Yoichi and Keiichi Yokobori, *Environment, Energy, and Economy: Strategies for Sustainability* (United Nations University Press, 1998).

Nakicenovic, Nebojsa, and Rob Swart (Eds.), *Special Report on Emissions Scenarios* (Cambridge University Press, 2000). https://www.ipcc.ch/report/emissions-scenarios/

Raupach, Michael R., *et al.*, "Global and regional drivers of accelerating CO2 emissions," PNAS **104**, 10288–10293 (2007) doi:10.1073/pnas.0700609104.

`emissions_factors` *Get emission factors for different energy sources*

Description

Get emission factors for different energy sources

Usage

```
emissions_factors(collapse_renewables = TRUE)
```

Arguments

`collapse_renewables`
: Combine hydroelectricity and other renewables into a single category.

Value

a tibble of values for emissions factors, in million metric tons of carbon dioxide per quad of energy.

See Also

regions

Examples

```
e_fac <- emissions_factors()
e_fac
```

`fuel_mix` *Mix of fuels contributing to primary energy supply for many countries and regions*

Description

A dataset containing the fuel mix of how many quads and what fraction of total primary energy supply comes from coal, gas, oil, nuclear, and renewable sources.

Usage

```
fuel_mix
```

Format

A tibble containing 948 rows and 7 variables

region Country or region name

region_code Three-letter country or region code

geography Geographic category: "nation", "region", or "world"

year The year

fuel The fuel: "Coal", "Natural Gas", "Oil", "Nuclear", "Hydro", and "Renewables"

quads The number of quads of that fuel consumed in the given country or region and year

frac The fraction of that country or region's total primary energy consumption from the fuel

Note

The 2023 data from the Energy Institute has inconsistencies in the fuel mix for Hong Kong and Sri Lanka: The percentages add up to 98.7% and 102.9%, respectively. The sums of energy in quads are off by -0.095 and +0.095 quads, respectively, from the total energy figure.

Source

https://www.energyinst.org/statistical-review/resources-and-data-downloads

See Also

regions, `get_fuel_mix()`

`generation_capacity` *Get power output from generation sources*

Description

Nameplate capacity and capacity factors for different electrical generation technologies. The average power supplied over a year is the nameplate capacity times the capacity factor.

Usage

```
generation_capacity()
```

Details

Data for fossil fuels comes from EIA

Value

a tibble of values for generation sources

fuel Energy source: Coal, Nuclear, Gas, Solar Thermal, Solar Photovoltaic, Onshore Wind, or Offshore Wind

description Text description of the power source

nameplate_capacity Maximum sustained power output, in megawatts

capacity_factor Capacity factor: the fraction of the nameplate capacity that the plant can provide, averaged over a typical year

References

Environmental Protection Agency (2018) "Electric Power Monthly," (October, 2018) https://www.eia.gov/electricity/monthly/archive/october2018.pdf, Table 6.7.A.

Pielke, Jr., Roger A., *The Climate Fix* (Basic Books, 2010).

Examples

```
gc <- generation_capacity()
gc
```

`get_fuel_mix` *Get fuel mix for one or more countries or regions*

Description

Get fuel mix for one or more countries or regions

Usage

```
get_fuel_mix(
  region_name,
  collapse_renewables = TRUE,
  quiet = FALSE,
  region_code = NULL
)
```

Arguments

`region_name`	A character vector with the names of one or more countries or regions to look up
`collapse_renewables`	Combine hydroelectricity and other renewables into a single category.
`quiet`	Suppress warnings if there is no data for that country or region.
`region_code`	Optional three-letter country or region codes to look up instead of the `region_name`

Value

A tibble of fuel mix for the countries or regions specified. That is, the number of quads of each fuel and the fraction of total primary energy coming from that fuel for each country or region:

region The name of the country or region

year The year reported

fuel The name of the fuel

quads The number of quads per year the country or region consumes

frac The fraction of the country's energy that comes from that fuel

Note

In the latest data from the Energy Institute, there are small discrepancies between the sums of energy for each fuel and the totals, in both `quads` and `frac`, for Hong Kong and Sri Lanka, as described in the documentation for fuel_mix.

See Also

regions, fuel_mix

Examples

```
get_fuel_mix("United States")
get_fuel_mix("World", collapse_renewables = FALSE)
get_fuel_mix(region_code = "LCN")
```

`get_kaya_data` *Get Kaya data for one or more countries or regions*

Description

Get Kaya data for one or more countries or regions

Usage

```
get_kaya_data(
  region_name,
  gdp = c("MER", "PPP"),
  quiet = FALSE,
  region_code = NULL
)
```

Arguments

`region_name`	The name of one or more countries or regions to look up
`gdp`	Use market exchange rates (`MER`) or purchasing power parity (`PPP`). Default is `MER`.
`quiet`	Suppress warnings if there is no such country or region.
`region_code`	Optional three-letter country or region codes to look up instead of the `region_name`

Details

Units for G, g, e, and ef depend on whether the data is requested in MER or PPP dollars: For MER, dollars are constant 2015 U.S. dollars. For PPP, dollars are constant 2017 international dollars.

```
_P_ and MER values for GDP and related quantities are available
from 1960 onward.

PPP values for GDP and related quantities are only available from
1990 onward.

Energy-related values (_E_, _F_, and derived quantities) are
available from 1965 onward.

Note that emissions (_F_, _f_, and _ef_) are reported as millions
of metric tons of carbon dioxide, not carbon.
```

Value

a tibble of Kaya identity data for the countries or regions specified:

region The name of the country or region

year The year

P Population, in billions

G Gross domestic product, in trillions of constant 2015 U.S. dollars.

E Total primary energy consumption, in quads

F CO2 emissions from fossil fuel consumption, in millions of metric tons

g Per-capita GDP, in thousands of dollars per person.

e Energy intensity of the economy, in quads per trillion dollars.

f Emissions intensity of the energy supply, in million metric tons per quad.

ef Emissions intensity of the economy, in metric tons per million dollars of GDP.

See Also

regions

Examples

```
get_kaya_data("Brazil")
get_kaya_data("United Kingdom", "PPP")
get_kaya_data(region_name = "United States")
get_kaya_data(region_code = "MYS")
```

---

get_top_down_trends *Get top-down trends for Kaya variables for one or more countries or regions, using projections from U.S. Energy Information Administration's International Energy Outlook report.*

Description

Get top-down trends for Kaya variables for one or more countries or regions, using projections from U.S. Energy Information Administration's International Energy Outlook report.

Usage

```
get_top_down_trends(region_name, quiet = FALSE, region_code = NULL)
```

Arguments

region_name	The name of one or more countries or regions to look up
quiet	Suppress warnings if there is no data for the specified countries or regions.
region_code	Optional three-letter country or region codes to look up instead of the region_name

Value

a tibble of trends for *P*, *G*, *E*, *F*, *g*, *e*, *f*, and *ef* for each country or region in percent per year.

See Also

regions

Examples

```
get_top_down_trends("Spain")
get_top_down_trends(region_code = "RUS")
```

`get_top_down_values` *Get top-down projections of Kaya variables for one or more countries or regions*

Description

Get top-down projections of Kaya variables for one or more countries or regions

Usage

```
get_top_down_values(region_name, quiet = FALSE, region_code = NULL)
```

Arguments

`region_name`	The name of a country or region to look up
`quiet`	Suppress warnings if there is no data for that country or region.
`region_code`	Optional three-letter country or region code to look up instead of the `region_name`

Value

a tibble of values for *P*, *G*, *E*, *F*, *g*, *e*, *f*, and *ef* for each country or region:

region The name of the country or region

P Population, in billions

G Gross domestic product, in trillions of constant 2015 U.S. dollars.

E Total primary energy consumption, in quads

F CO2 emissions from fossil fuel consumption, in millions of metric tons

g Per-capita GDP, in thousands of constant 2015 U.S. dollars per person.

e Energy intensity of the economy, in quads per trillion dollars.

f Emissions intensity of the energy supply, in million metric tons per quad.

ef Emissions intensity of the economy, in metric tons per million dollars of GDP.

See Also

regions

Examples

```
get_top_down_values("New Zealand")
get_top_down_values("OECD")
get_top_down_values(region_code = "PAK")
```

---

kaya_data *Kaya identity data for many countries and regions*

Description

A dataset containing Kaya identity parameters P, G, E, F, g, e, f, and ef for many countries

Usage

```
kaya_data
```

Format

A tibble containing 5292 rows and 14 variables:

region Country or region name

region_code Three-letter country or region code

geography Geographic category: "nation", "region", or "world"

year The year

P Population, in billions

G Gross domestic product, in trillions of constant 2015 U.S. dollars.

E Total primary energy consumption, in quads

F CO2 emissions from fossil fuel consumption, in millions of tons

g Per-capita GDP, in thousands of constant 2015 U.S. dollars per person.

e Energy intensity of the economy, in quads per trillion dollars.

f Emissions intensity of the energy supply, in million metric tons per quad.

ef Emissions intensity of the economy, in metric tons per million dollars of GDP.

G_ppp Gross domestic product adjusted for purchasing power parity, in trillions of constant 2017 international dollars

G_mer Gross domestic product at market-exchange-rate, in trillions of constant 2015 U.S. dollars

Source

https://data.worldbank.org/indicator/SP.POP.TOTL, https://data.worldbank.org/indicator/NY.GDP.MKTP.KD, and https://www.energyinst.org/statistical-review

See Also

regions, get_kaya_data()

`kaya_region_list` *Get a list of countries in the Kaya data*

Description

Get a list of countries in the Kaya data

Usage

```
kaya_region_list()
```

Value

a vector of country and region names

See Also

regions

`megawatts_per_quad` *The number of megawatts it takes to replace a quad.*

Description

The number of megawatts of average power output over a year to produce one quad of energy

Usage

```
megawatts_per_quad()
```

Value

The number of megawatts equivalent to one quad per year.

Examples

```
mwe <- megawatts_per_quad()
mwe
```

plot_fuel_mix *Plot fuel mix*

Description

Plot fuel mix

Usage

```
plot_fuel_mix(
  fuel_mix,
  collapse_renewables = TRUE,
  title = NULL,
  colors = NULL,
  font_size = 20
)
```

Arguments

fuel_mix	A tibble with the mixture of fuels for one or more countries or regions: **region** The name of the country or region **fuel** The name of the fuel **quads** The number of quads per year the country or region consumes **frac** The fraction of the country's energy that comes from that fuel
collapse_renewables	Combine hydroelectricity and other renewables into a single category.
title	Include a title on the plot. If `title` is `NULL` (default) or `TRUE`, a default title is created from the names of the regions in `fuel_mix`. If `title` is a character string, that string is used. If `title` is `FALSE`, the plot is produced with no title.
colors	A named vector with the colors to use for `Coal`, `Oil`, `Natural Gas`, `Nuclear`, `Hydro`, and `Renewables`.
font_size	The base font size.

Value

A plot object.

Examples

```
usa_fuel <- get_fuel_mix("United States", collapse_renewables = FALSE)
plot_fuel_mix(usa_fuel)
plot_fuel_mix(usa_fuel, collapse_renewables = FALSE)
plot_fuel_mix(usa_fuel, collapse_renewables = FALSE,
              colors = c(Coal = "black", "Natural Gas" = "gray60",
                         Oil = "gray30", Nuclear = "forestgreen",
                         Hydro = "royalblue", Renewables="palegreen"))
```

plot_kaya *Plot Kaya-identity variable*

Description

Plot Kaya-identity variable

Usage

```
plot_kaya(
  kaya_data,
  variable,
  start_year = NA,
  stop_year = NA,
  y_lab = NULL,
  log_scale = FALSE,
  trend_line = FALSE,
  points = TRUE,
  font_size = 20,
  colors = NULL,
  pre_color = NULL,
  post_color = NULL,
  in_range_color = NULL,
  trend_color = NULL,
  line_sizes = NULL,
  pre_line_size = NULL,
  post_line_size = NULL,
  in_range_line_size = NULL,
  trend_line_size = NULL,
  point_sizes = NULL,
  pre_point_size = NULL,
  post_point_size = NULL,
  in_range_point_size = NULL
)
```

Arguments

kaya_data	A tibble with Kaya-identity data
variable	The name of the variable to plot (character)
start_year	The year to start highlighting the data (should correspond to the beginning of the trend calculation). Set to NULL to turn off highlighting.
stop_year	The year to stop highlighting the data (should correspond to the beginning of the trend calculation). Set to NULL to turn off highlighting.
y_lab	Optional label for the y-axis
log_scale	Use log scale for y axis

trend_line	Include a trend line
points	Plot points in addition to the line.
font_size	Base size of the font for axis labels and titles.
colors	Named vector of colors to use for the plot. Elements should include PRE, POST, IN-RANGE, and TREND, which respectively give the colors for the portion of the plot before start_year, after stop_year, between start_year and stop_year, and the trend line.
pre_color	Override default color for the portion of the chart before start_year.
post_color	Override default color for the portion of the chart after stop_year.
in_range_color	Override default color for the portion of the chart between start_year and stop_year.
trend_color	Override default color for the trend line.
line_sizes	Named vector of sizes to use for the lines in the plot. Elements should include PRE, POST, IN-RANGE, and TREND, which respectively give the sizes for lines in the portion of the plot before start_year, after stop_year, between start_year and stop_year, and the trend line.
pre_line_size	Override default line size for the portion of the chart before start_year.
post_line_size	Override default line size for the portion of the chart after stop_year.
in_range_line_size	Override default line size for the portion of the chart between start_year and stop_year.
trend_line_size	Override default size for the trend line.
point_sizes	Named vector of sizes to use for the points in the plot. Elements should include PRE, POST, and IN-RANGE, which respectively give the sizes for points in the portion of the plot before start_year, after stop_year, and between start_year and stop_year.
pre_point_size	Override default point size for the portion of the chart before start_year.
post_point_size	Override default point size for the portion of the chart after stop_year.
in_range_point_size	Override default point size for the portion of the chart between start_year and stop_year.

Value

A plot object.

Examples

```
china <- get_kaya_data("China")
plot_kaya(china, "F", 2001, 2011)
## Not run:
uk <- get_kaya_data("United Kingdom")
plot_kaya(uk, "e", log_scale = TRUE, trend_line = TRUE)
```

```
plot_kaya(uk, "e", log_scale = TRUE, trend_line = TRUE,
          start_year = 1970, stop_year = 2000,
          colors = c(PRE="limegreen", POST="darkgreen",
                     "IN-RANGE" = "cadetblue", TREND="orange"),
          line_sizes = c(PRE=0.5, POST=0.5, "IN-RANGE"=1, TREND=1.5),
          point_sizes = c(PRE=2, POST=2, "IN-RANGE"=3))
plot_kaya(uk, "e", log_scale = TRUE, trend_line = TRUE,
          start_year = 1970, stop_year = 2000,
          pre_color = "limegreen", post_color = "limegreen",
          trend_color = "magenta",
          pre_line_size = 0.5, post_line_size = 0.5,
          trend_line_size = 1.5,
          pre_point_size = 2, post_point_size = 2, in_range_point_size = 3)

## End(Not run)

world <- get_kaya_data("World")
plot_kaya(world, "g", 1982, log_scale = TRUE, trend_line = TRUE)
```

`project_top_down` *Get top-down projections of Kaya variables for one or more countries or regions for a given year*

Description

Get top-down projections of Kaya variables for one or more countries or regions for a given year

Usage

```
project_top_down(region_name, year, quiet = FALSE, region_code = NULL)
```

Arguments

`region_name`	The name of a country or region to look up
`year`	The year to project to
`quiet`	Suppress warnings if there is no data for that country or region.
`region_code`	Optional three-letter country or region code to look up instead of the `region_name`

Value

a tibble of values for P, G, E, F, g, e, f, and ef for each country or region:

region The name of the country or region

year The year

P Population, in billions

G Gross domestic product, in trillions of constant 2015 U.S. dollars.

E Total primary energy consumption, in quads

F CO2 emissions from fossil fuel consumption, in millions of metric tons

g Per-capita GDP, in thousands of constant 2015 U.S. dollars per person.

e Energy intensity of the economy, in quads per trillion dollars.

f Emissions intensity of the energy supply, in million metric tons per quad.

ef Emissions intensity of the economy, in metric tons per million dollars of GDP.

See Also

regions

Examples

```
project_top_down("China", 2037)
project_top_down(region_code = "VNM", year = 2043)
```

`regions` *Aggregate regional data*

Description

Problems with aggregate regional data

Details

The World Bank is missing GDP data for a number of nations, such as Syria and Taiwan. Because of this and the incommensurability between the regions used for aggregate statistics in the World Bank and Energy Institute data, aggregate regional data (e.g., for the Middle East and Africa) should be treated with caution. This problem does not hold for individual nations, where missing data appears as `NA` values.

`td_trends` *Top-down projections of trends in Kaya variables for many countries and regions*

Description

A dataset containing top-down projections of trends in P, G, and E, from the EIA's International Energy Outlook 2017.

Usage

```
td_trends
```

Format

A tibble containing 226 rows and 11 variables

region Country or region name

region_code Three-letter country or region code

geography Geographic category: "nation", "region", or "world"

P Trend in population, in fraction per year

G Trend in gross domestic product, in fraction per year

E Trend in total primary energy consumption, in fraction per year

F Trend in CO2 emissions, in fraction per year

g Trend in per-capita GDP, in fraction per year

e Trend in energy intensity of the economy, in fraction per year

f Trend in emissions intensity of the energy supply, in fraction per year

ef Trend in emissions intensity of the economy, in fraction per year

Source

https://www.eia.gov/outlooks/archive/ieo17/

See Also

regions, get_top_down_trends()

---

td_values *Top-down projections of future Kaya variables for many countries and regions*

Description

A dataset containing top-down projections of P, G, and E, from the EIA's International Energy Outlook 2017.

Usage

```
td_values
```

Format

A tibble containing 640 rows and 12 variables

region Country or region name

region_code Three-letter country or region code

geography Geographic category: "nation", "region", or "world"

year The year

P Population, in billions

G Gross domestic product, in trillions of constant 2015 U.S. dollars

E Total primary energy consumption, in quads

F Total CO2 emissions, in millions of metric tons

g Per-capita GDP, in thousands of constant 2015 U.S. dollars per person.

e Energy intensity of the economy, in quads per trillion dollars.

f Emissions intensity of the energy supply, in million metric tons per quad.

ef Emissions intensity of the economy, in metric tons per million dollars of GDP.

Source

https://www.eia.gov/outlooks/archive/ieo17/

See Also

regions, get_top_down_values()

Index