Package 'epidata'

April 11, 2019

Type Package

Title Tools to Retrieve Economic Policy Institute Data Library Extracts

Version 0.3.0

Date 2019-04-11

Maintainer Bob Rudis <bob@rud.is>

Encoding UTF-8

Description The Economic Policy Institute (<http://www.epi.org/>) provides researchers, media, and the public with easily accessible, up-to-date, and comprehensive historical data on the American labor force. It is compiled from Economic Policy Institute analysis of government data sources. Use it to research wages, inequality, and other economic indicators over time and among demographic groups. Data is usually updated monthly.

URL https://gitlab.com/hrbrmstr/epidata

BugReports https://gitlab.com/hrbrmstr/epidata/issues

License AGPL

Suggests testthat

Depends R (>= 3.2.0)

Imports purrr, httr, jsonlite, dplyr, rvest, xml2, tidyr, readr, stringi

RoxygenNote 6.1.1

NeedsCompilation no

Author Bob Rudis [aut, cre]

Repository CRAN

Date/Publication 2019-04-11 18:32:40 UTC

R topics documented:

`get_annual_wages_and_work_hours` *Retreive CPS ASEC Annual Wages and Work Hours*

Description

Annual, weekly, and hourly wages and work hours show the average wages and work hours of wage and salary workers using data from the CPS ASEC (also known as the March CPS). Note that this data is not directly comparable to the CPS ORG data in median/average hourly wage.

Usage

```
get_annual_wages_and_work_hours()
```

Value

`tbl_df`

Note

CPS ASEC | Murphy and Welch (1989)

References

Economic Policy Institute Data Library

Examples

```
get_annual_wages_and_work_hours()
```

`get_annual_wages_by_wage_group`
Annual wages by wage group

Description

Return the average annual salaries for select wage groups, with particular focus on the highest wage earners. Note that this data is not directly comparable to wage deciles/percentiles.

Usage

```
get_annual_wages_by_wage_group()
```

Details

Wages are in 2017 dollars. Population sample: All workers.

The average annual wages by wage group are taken from a 2010 article by Wojciech Kopczuk, Emmanuel Saez, and Jae Song. To extend this series, data for 2006 through 2017 are extrapolated from 2004 data using changes in wage shares computed from Social Security Administration wage statistics. We employ the midpoint of the bracket to compute total wage income in each bracket and sum all brackets. We then use interpolation to derive percentile cutoffs building from the bottom up to obtain the 0–90th percentile bracket and then estimate the remaining categories. This allows us to estimate the wage shares for upper wage groups. We use these wage shares computed for 2004 and later years to extend the Kopczuk, Saez, and Song series by adding the changes in share between 2004 and the relevant year to their series. To obtain absolute wage trends we use the SSA data on the total wage pool and employment and compute the real wage per worker (based on t heir share of wages and employment) in the different groups in 2017 dollars. For a detailed explanation, see the methodology for annual wages and hours.

Value

`tbl_df` with data filtered by the selected criteria.

Note

Data source: SSA | Kopczuk, Saez, and Song (2010)

References

Economic Policy Institute Data Library

Examples

```
get_annual_wages_by_wage_group()
```

`get_black_white_wage_gap` *Retreive the percent by which hourly wages of black workers are less than hourly wages of white workers*

Description

The black-white wage gap is the percent by which hourly wages of black workers are less than hourly wages of white workers. It is also often expressed as a wage ratio (black workers' share of white workers' wages) by subtracting the gap from 100 percent.

Usage

```
get_black_white_wage_gap(by = NULL)
```

Arguments

by	`NULL` or g for a parition by gender

Details

- A median black-white wage gap of 26.2 percent means that a typical black worker is paid 26.2 percent less per hour than a typical white worker.
- An average black-white wage gap of 26.6 percent means that on average black workers are paid 26.6 percent less per hour than white workers.
- A regression-based black-white wage gap of 15.2 percent means that on average black workers are paid 15.2 percent less per hour than white workers, all else held equal (controlling for gender, race and ethnicity, education, experience, and geographic location).

Value

`tbl_df` with data filtered by the selected criteria.

References

Economic Policy Institute Data Library

Examples

```
get_black_white_wage_gap()

get_black_white_wage_gap("g")
```

get_college_wage_premium

Retreive the percent by which hourly wages of college graduates exceed those of otherwise equivalent high school graduates

Description

A regression-based college wage premium of 56.1 percent means that on average workers with a college degree are paid 56.1 percent more per hour than workers whose highest education credential is a high school diploma, all else held equal (controlling for gender, race and ethnicity, education, experience, and geographic location).

Usage

```
get_college_wage_premium(by = NULL)
```

Arguments

by	NULL or g for a parition by gender

Value

tbl_df with data filtered by the selected criteria.

References

Economic Policy Institute Data Library

Examples

```
get_college_wage_premium()

get_college_wage_premium("g")
```

get_compensation_wages_and_benefits

Compensation, wages, and benefits

Description

Return the nonwage payments, referred to as fringe benefits, and wages. Compensation includes employer payments for health insurance, pensions, and payroll taxes (primarily payments toward Social Security and unemployment insurance).

Usage

```
get_compensation_wages_and_benefits()
```

Details

Wages are in 2016 dollars. Wage and salary workers (NIPA) | Private-sector workers (ECEC)

Value

`tbl_df` with data filtered by the selected criteria.

Note

Data source: NIPA | ECEC

References

Economic Policy Institute Data Library

Examples

```
get_compensation_wages_and_benefits()
```

`get_employment_to_population_ratio`

Retreive the share of the civilian noninstitutional population that is employed

---

Description

Retreive the share of the civilian noninstitutional population that is employed

Usage

```
get_employment_to_population_ratio(by = NULL)
```

Arguments

`by`	`NULL` or character string with any combination of `g` (Gender), `r` (Race), `a` (Age), `e` (Education). i.e. if you want to retrieve unemployment data by gender, race and education, you would set this parameter to "`gre`".

Value

`tbl_df` with data filtered by the selected criteria.

References

Economic Policy Institute Data Library

Examples

```
get_employment_to_population_ratio()

get_employment_to_population_ratio("r")

get_employment_to_population_ratio("grae")
```

`get_gender_wage_gap` *Retreive the percent by which hourly wages of female workers are less than hourly wages of male workers*

Description

The gender wage gap is the percent by which hourly wages of female workers are less than hourly wages of male workers. It is also often expressed as a wage ratio (women's share of men's wages) by subtracting the gap from 100 percent.

Usage

```
get_gender_wage_gap(by = NULL)
```

Arguments

by	`NULL` or `r` for a parition by race

Details

- A median gender wage gap of 17.3 percent means that a typical woman is paid 17.3 percent less per hour than a typical man.
- An average gender wage gap of 19.7 percent means that on average women are paid 19.7 percent less per hour than men.
- A regression-based gender wage gap of 21.7 percent means that on average women are paid 21.7 percent less per hour than men, all else held equal (controlling for gender, race and ethnicity, education, experience, and geographic location).

Value

`tbl_df` with data filtered by the selected criteria.

References

Economic Policy Institute Data Library

Examples

```
get_gender_wage_gap()

get_gender_wage_gap("r")
```

`get_health_insurance_coverage` *Retreive Health Insurance Coverage*

Description

Employer-sponsored health insurance (ESI) coverage shows the share of workers who received health insurance from their own job for which their employer paid for at least some of their health insurance coverage.

Usage

```
get_health_insurance_coverage(by = NULL)
```

Arguments

by	`NULL` or character string with any combination of `g` (Gender), `r` (Race), `e` (Education), `d` (Percentile), `l` (Entry-level) i.e. if you want to retrieve unemployment data by gender and race, you would set this parameter to "`gr`".

Details

Population sample: Private-sector workers age 18–64 & at least 20 hours/week and 26 weeks/year

Value

`tbl_df` with data filtered by the selected criteria.

Note

Data source: CPS ASEC

References

Economic Policy Institute Data Library

Examples

```
get_health_insurance_coverage()

get_health_insurance_coverage("r")

get_health_insurance_coverage("gr")
```

`get_hispanic_white_wage_gap` *Retreive the percent by which hourly wages of Hispanic workers are less than hourly wages of white workers*

Description

The Hispanic-white wage gap is the percent by which hourly wages of Hispanic workers are less than hourly wages of white workers. It is also often expressed as a wage ratio (Hispanic workers' share of white workers' wages) by subtracting the gap from 100 percent.

Usage

```
get_hispanic_white_wage_gap(by = NULL)
```

Arguments

by	NULL or g for a parition by gender

Details

- A median Hispanic-white wage gap of 29.6 percent means that a typical Hispanic worker is paid 29.6 percent less per hour than a typical white worker.
- An average Hispanic-white wage gap of 30.1 percent means that on average Hispanic workers are paid 30.1 percent less per hour than white workers.
- A regression-based Hispanic-white wage gap of 11.1 percent means that on average Hispanic workers are paid 11.1 percent less per hour than white workers, all else held equal (controlling for gender, race and ethnicity, education, experience, and geographic location).

Value

`tbl_df` with data filtered by the selected criteria.

References

Economic Policy Institute Data Library

Examples

```
get_hispanic_white_wage_gap()

get_hispanic_white_wage_gap("g")
```

`get_labor_force_participation_rate`

Retreive the share of the civilian noninstitutional population that is in the labor force

---

Description

(i.e., working or looking for work)

Usage

```
get_labor_force_participation_rate(by = NULL)
```

Arguments

by	NULL or character string with any combination of g (Gender), r (Race), a (Age), e (Education). i.e. if you want to retrieve unemployment data by gender, race and education, you would set this parameter to "gre".

Value

tbl_df with data filtered by the selected criteria.

References

Economic Policy Institute Data Library

Examples

```
get_labor_force_participation_rate()

get_labor_force_participation_rate("r")

get_labor_force_participation_rate("grae")
```

`get_long_term_unemployment`

Retreive the share of the labor force that has been unemployed for six months or longer

---

Description

Retreive the share of the labor force that has been unemployed for six months or longer

Usage

```
get_long_term_unemployment(by = NULL)
```

Arguments

by	NULL or character string with any combination of g (Gender), r (Race), a (Age), e (Education). i.e. if you want to retrieve unemployment data by gender, race and education, you would set this parameter to "gre".

Value

tbl_df with data filtered by the selected criteria.

References

Economic Policy Institute Data Library

Examples

```
get_long_term_unemployment()

get_long_term_unemployment("r")

get_long_term_unemployment("grae")
```

---

get_median_and_mean_wages

Retreive the hourly wage in the middle of the wage distribution

---

Description

The median wage is the hourly wage in the middle of the wage distribution; 50 percent of wage earners earn less and 50 percent earn more. The average wage is the arithmetic mean of hourly wages; or, the sum of all workers' hourly wages divided by the number of workers.

Usage

```
get_median_and_mean_wages(by = NULL)
```

Arguments

by	NULL or character string with any combination of g (Gender), r (Race), e (Education), d (Percentile), l (Entry-level) i.e. if you want to wage data by gender and race, you would set this parameter to "gr".

Value

tbl_df with data filtered by the selected criteria.

References

Economic Policy Institute Data Library

Examples

```
get_median_and_mean_wages()

get_median_and_mean_wages("r")

get_median_and_mean_wages("gr")
```

`get_minimum_wage` *Minimum wage*

Description

Return tthe hourly minimum wage set by federal law. The real minimum wage is the federal hourly minimum wage adjusted for inflation.

Usage

```
get_minimum_wage()
```

Details

Wages are in 2016 dollars, excluding the nominal federal minimum wage. Share of average wages based on the average wages of production and nonsupervisory workers. For state minimum wages, see EPI's minimum wage tracker.

Population sample: Production and nonsupervisory workers (average wages)

Value

`tbl_df` with data filtered by the selected criteria.

Note

Data source: U.S. Department of Labor Wage and Hour Division | CES

References

Economic Policy Institute Data Library

Examples

```
get_minimum_wage()
```

`get_non_high_school_wage_penalty` *Retreive the percent by which hourly wages of workers without a high school diploma (or equivalent) are less than wages of otherwise equivalent workers who have graduated from high school*

Description

A regression-based non-high school wage penalty of 21.8 percent means that on average workers without a high school diploma are paid 21.8 percent less per hour than workers with a high school diploma, all else held equal (controlling for gender, race and ethnicity, education, experience, and geographic location).

Usage

```
get_non_high_school_wage_penalty(by = NULL)
```

Arguments

by	`NULL` or g for a parition by gender

Value

`tbl_df` with data filtered by the selected criteria.

References

Economic Policy Institute Data Library

Examples

```
## Not run:
get_non_high_school_wage_penalty()

get_non_high_school_wage_penalty("g")

## End(Not run)
```

`get_pension_coverage` *Retreive Pension Coverage*

Description

Employer-provided pension coverage shows the share of workers included in an employer-provided plan for which the employer paid for at least some of their pension coverage.

Usage

```
get_pension_coverage(by = NULL)
```

Arguments

by	`NULL` or character string with any combination of `g` (Gender), `r` (Race), `e` (Education), `d` (Percentile), `l` (Entry-level) i.e. if you want to retrieve pension data by gender and race, you would set this parameter to `"gr"`.

Details

Population sample: Private-sector workers age 18–64 & at least 20 hours/week and 26 weeks/year

Value

`tbl_df` with data filtered by the selected criteria.

Note

Data source: CPS ASEC

References

Economic Policy Institute Data Library

Examples

```
get_health_insurance_coverage()

get_health_insurance_coverage("r")

get_health_insurance_coverage("gr")
```

`get_poverty_level_wages` *Poverty-level wages*

Description

Return the share of workers earning equal to or less than the poverty-level wage, or the hourly wage that a full-time, year-round worker must earn to sustain a family of four with two children at the official poverty threshold.

Usage

```
get_poverty_level_wages(by = NULL)
```

Arguments

by	NULL or character string with any combination of g (Gender) or r (Race), i.e. if you want to retrieve unemployment data by gender and race, you would set this parameter to "gr".

Details

Population sample: Wage and salary workers age 18–64. Data source: CPS ORG | Census Bureau (poverty threshold)

Value

tbl_df with data filtered by the selected criteria.

References

Economic Policy Institute Data Library

Examples

```
## Not run:
get_poverty_level_wages()

get_poverty_level_wages("r")

get_poverty_level_wages("gr")

## End(Not run)
```

`get_productivity_and_hourly_compensation` *Retreive Productivity and hourly compensation*

Description

Productivity is how much workers produce per hour, or the growth of output of goods and services minus depreciation per hour worked. Compensation is made up of both nonwage payments and wages.

Usage

```
get_productivity_and_hourly_compensation(by = NULL)
```

Arguments

by	NULL or character string of g (Gender)

Details

Wages are in 2015 dollars. Median compensation is calculated using hourly wage medians from the CPS ORG and compensation from NIPA.

Population sample: All workers & Production and nonsupervisory workers

Value

`tbl_df` with data filtered by the selected criteria.

Note

Data source: NIPA (compensation) | BLS Productivity Data

References

Economic Policy Institute Data Library

Examples

```
get_productivity_and_hourly_compensation()

get_productivity_and_hourly_compensation("g")
```

get_underemployment *Retreive the share of the labor force that is "underemployed"*

Description

Underemployment is the share of the labor force that either 1) is unemployed, 2) is working part time but wants and is available to work full time (an "involuntary" part timer), or 3) wants and is available to work and has looked for work in the last year but has given up actively seeking work in the last four weeks ("marginally attached" worker).

Usage

```
get_underemployment(by = NULL)
```

Arguments

by	NULL or character string with any combination of g (Gender), r (Race), a (Age), e (Education). i.e. if you want to retrieve unemployment data by gender, race and education, you would set this parameter to "gre".

Value

tbl_df with data filtered by the selected criteria.

References

Economic Policy Institute Data Library

Examples

```
get_underemployment()

get_underemployment("r")

get_underemployment("grae")
```

get_unemployment *Retreive the share of the labor force without a job*

Description

Retreive the share of the labor force without a job

Usage

```
get_unemployment(by = NULL)
```

Arguments

by	NULL or character string with any combination of g (Gender), r (Race), a (Age), e (Education). i.e. if you want to retrieve unemployment data by gender, race and education, you would set this parameter to "gre".

Value

tbl_df with data filtered by the selected criteria.

Note

See get_unemployment_by_state() for information on retrieving unemployment by state+race.

References

Economic Policy Institute Data Library

Examples

```
get_unemployment()

get_unemployment("r")

get_unemployment("grae")
```

---

get_unemployment_by_state

Retreive the share of the labor force without a job (by state)

Description

Retreive the share of the labor force without a job (by state)

Usage

```
get_unemployment_by_state(by = NULL)
```

Arguments

by	NULL or r for a partition by race.

Value

tbl_df with data filtered by the selected criteria.

Note

See get_unemployment() for other unemployment extracts..

References

Economic Policy Institute Data Library

Examples

```
get_unemployment_by_state()

get_unemployment_by_state("r")
```

`get_union_coverage` *Retreive Union Coverage*

Description

The union coverage rate shows the percentage of the workforce covered by a collective bargaining agreement.

Usage

```
get_union_coverage()
```

Value

`tbl_df`

Note

Data source: CPS ORG | Hirsch and Macpherson (2003)

References

Economic Policy Institute Data Library

Examples

```
get_union_coverage()
```

get_wages_by_education

Retreive the average hourly wages of workers disaggregated by the highest level of education attained

Description

Wages by education are the average hourly wages of workers disaggregated by the highest level of education attained. Employment shares provide the distribution of educational attainment for workers of each gender, racial, and ethnic group as a share of total employed for each group.

Usage

```
get_wages_by_education(by = NULL)
```

Arguments

by	NULL or character string with any combination of g (Gender) or r (Race), i.e. if you want to retrieve unemployment data by gender and race, you would set this parameter to "gr".

Value

tbl_df with data filtered by the selected criteria.

References

Economic Policy Institute Data Library

Examples

```
get_wages_by_education()

get_wages_by_education("r")

get_wages_by_education("gr")
```

get_wages_by_percentile

Retreive wages at ten distinct points in the wage distribution

Description

Wage percentiles are wages at ten distinct points in the wage distribution: deciles and the 95th percentile. The 95–50 and 50–10 wage ratios show how much greater wages are at the top than the middle, and at the middle than the bottom, respectively.

Usage

```
get_wages_by_percentile(by = NULL)
```

Arguments

by	NULL or character string with any combination of g (Gender) or r (Race), i.e. if you want to retrieve unemployment data by gender and race, you would set this parameter to "gr".

Value

tbl_df with data filtered by the selected criteria.

References

Economic Policy Institute Data Library

Examples

```
get_wages_by_percentile()

get_wages_by_percentile("r")

get_wages_by_percentile("gr")
```

---

get_wage_decomposition

Retreive Wage Decomposition

Description

Wage inequality data shows the overall wage inequality and the within-group and between-group wage inequality over time. These measures allow an examination of how much of the change in overall wage inequality in particular periods was due to changes in within-group and between-group wage inequality.

Usage

```
get_wage_decomposition(by = NULL)
```

Arguments

by	NULL or character string of g (Gender)

Details

Population sample: Wage and salary workers age 18–64

Value

tbl_df with data filtered by the selected criteria.

Note

Data source: CPS ORG

References

Economic Policy Institute Data Library

Examples

```
get_wages_by_percentile()

get_wages_by_percentile("g")
```

---

get_wage_ratios *Retreive the level of inequality within the hourly wage distribution.*

Description

The 95–50 and 50–10 wage ratios are representations of the level of inequality within the hourly wage distribution. The larger the ratio, the greater the gap between the top and the middle or the middle and the bottom of the wage distribution.

Usage

```
get_wage_ratios(by = NULL)
```

Arguments

by	NULL or character string with any combination of g (Gender) or r (Race), i.e. if you want to retrieve unemployment data by gender and race, you would set this parameter to "gr".

Details

- A 50–10 wage ratio of 1.91 means that workers at the 50th percentile of the wage distribution are paid 1.91 times more per hour than the workers at the 10th percentile.
- A 95–50 wage ratio of 3.28 means that workers at the 95th percentile of the wage distribution are paid 3.28 times more per hour than the workers at the 50th percentile.

Value

tbl_df with data filtered by the selected criteria.

References

Economic Policy Institute Data Library

Examples

```
## Not run:
get_wage_ratios()

get_wage_ratios("r")

get_wage_ratios("gr")

## End(Not run)
```

Index