

# EC 502 HW3 R Notebook

## Import tidyverse

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
## v dplyr      1.1.4      v readr      2.1.5
## v forcats    1.0.0      v stringr   1.5.1
## v ggplot2    3.4.4      v tibble    3.2.1
## v lubridate  1.9.3      v tidyr     1.3.1
## v purrr      1.0.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

## Load the dataset:

```
country_data = read.csv("cross_country_data.csv")
head(country_data, 10)
```

##	Country	Year	SavingsRate	RealGDP	LaborForce	HumanCapital	PhysicalCapital
## 1	Albania	1980	0.3934749	9981.908	0.9249071	2.395581	42834.52
## 2	Albania	1981	0.4208891	10555.427	0.9514943	2.445612	45231.09
## 3	Albania	1982	0.4482948	10866.670	0.9771618	2.496687	47943.00
## 4	Albania	1983	0.4195175	10986.740	1.0006725	2.548830	50831.43
## 5	Albania	1984	0.3768510	10849.229	1.0246358	2.588237	53751.48
## 6	Albania	1985	0.3867719	11042.415	1.0539739	2.624506	56010.12
## 7	Albania	1986	0.3833166	11660.205	1.0901078	2.648263	58482.48
## 8	Albania	1987	0.3503911	11568.308	1.1281416	2.672236	61068.25
## 9	Albania	1988	0.3539001	11403.967	1.1635048	2.696425	63407.10
## 10	Albania	1989	0.3725055	12526.185	1.1956922	2.720833	66093.03

## Calculate (real) GDP per worker :

```
country_data <- transform(country_data, rgdp_per_worker = RealGDP / LaborForce)
head(country_data[c('Country', 'RealGDP', 'LaborForce', 'rgdp_per_worker')], 10)
```

##	Country	RealGDP	LaborForce	rgdp_per_worker
## 1	Albania	9981.908	0.9249071	10792.336

```
## 2 Albania 10555.427 0.9514943 11093.526
## 3 Albania 10866.670 0.9771618 11120.646
## 4 Albania 10986.740 1.0006725 10979.357
## 5 Albania 10849.229 1.0246358 10588.376
## 6 Albania 11042.415 1.0539739 10476.934
## 7 Albania 11660.205 1.0901078 10696.378
## 8 Albania 11568.308 1.1281416 10254.304
## 9 Albania 11403.967 1.1635048 9801.392
## 10 Albania 12526.185 1.1956922 10476.095
```

Average rGDP per worker and average savings rate:

```
head(country_data %>%
  group_by(Country) %>%
  summarise_at(vars(rgdp_per_worker, SavingsRate), list(mean = mean)), 10)
```

```
## # A tibble: 10 x 3
##   Country    rgdp_per_worker_mean SavingsRate_mean
##   <chr>          <dbl>          <dbl>
## 1 Albania      12856.          0.279
## 2 Argentina    26441.          0.185
## 3 Australia    63532.          0.287
## 4 Austria      61305.          0.280
## 5 Bahrain      65812.          0.259
## 6 Bangladesh   3565.           0.166
## 7 Barbados     50524.          0.0914
## 8 Belgium      68129.          0.270
## 9 Belize       15970.          0.145
## 10 Benin       3635.           0.108
```

Average growth rate of the labor force and average growth rate of GDP per worker (from 1981 to 2010):

$t = 1981$

$T = 2010 - 1981 = 29$

formula:  $1/29 * \ln(\text{val at 2010} / \text{val at 1981})$

(log in R has base e)

```
country_data_2010_1980 <- country_data %>%
  filter(Year %in% c(1980, 2010)) %>%
  select(Country, Year, LaborForce, rgdp_per_worker) %>%
  pivot_wider(names_from = Year,
    values_from = c(rgdp_per_worker, LaborForce),
    names_sep = "_",
    names_prefix = "year_") %>%
  transform(avg_growth_labor = (1/29) * log(LaborForce_year_2010 / LaborForce_year_1980)) %>%
  transform(avg_gdp_per_worker = (1/29) * log(rgdp_per_worker_year_2010 / rgdp_per_worker_year_1980))

# only show avg growth rates
head(country_data_2010_1980 %>% select(Country, avg_growth_labor, avg_gdp_per_worker), 10)
```

```
##      Country avg_growth_labor avg_gdp_per_worker
## 1   Albania  -0.000296065      0.0279996634
## 2   Argentina  0.013595292      0.0114475755
## 3   Australia  0.020844982      0.0119543025
## 4   Austria   0.009453424      0.0125808953
## 5   Bahrain   0.057890955     -0.0230461528
## 6   Bangladesh 0.025913951      0.0225090252
## 7   Barbados  0.010562211     -0.0006532531
## 8   Belgium   0.006871757      0.0123780049
## 9   Belize    0.033855566      0.0136094263
## 10  Benin     0.032257195      0.0085833431
```

## Question (1) Summary Stats

a) number of countries

```
country_data %>% summarise(num_countries = n_distinct(Country))
```

```
##   num_countries
## 1             118
```

b) mean (across countries) of rgdp per worker 1981, rgdp per worker 2010, savings rate, growth of labor force, growth rate of GDP per worker

```
country_data_2010_1980 %>%
  summarise_at(vars(rgdp_per_worker_year_1980, rgdp_per_worker_year_2010, avg_growth_labor, avg_gdp_per_
  pivot_longer(cols=everything(), names_to = "variable", values_to = "value"))
```

```
## # A tibble: 4 x 2
##   variable                                value
##   <chr>                                <dbl>
## 1 rgdp_per_worker_year_1980_mean 28519.
## 2 rgdp_per_worker_year_2010_mean 34551.
## 3 avg_growth_labor_mean          0.0233
## 4 avg_gdp_per_worker_mean        0.0114
```

Savings rate (s):

```
country_data %>%
  summarise_at(vars(SavingsRate), list(mean=mean))
```

```
##      mean
## 1 0.2003188
```

c) standard deviation of vals above

```
country_data_2010_1980 %>%
  summarise_at(vars(rgdp_per_worker_year_1980, rgdp_per_worker_year_2010, avg_growth_labor, avg_gdp_per_worker_year_2010),
    pivot_longer(cols=everything(), names_to = "variable", values_to = "value"))
```

```
## # A tibble: 4 x 2
##   variable      value
##   <chr>      <dbl>
## 1 rgdp_per_worker_year_1980_sd 46105.
## 2 rgdp_per_worker_year_2010_sd 33102.
## 3 avg_growth_labor_sd         0.0137
## 4 avg_gdp_per_worker_sd         0.0190
```

Savings rate (s):

```
country_data %>%
  summarise_at(vars(SavingsRate), list(sd=sd)) %>%
  pivot_longer(cols=everything(), names_to = "variable", values_to = "value")
```

```
## # A tibble: 1 x 2
##   variable value
##   <chr>      <dbl>
## 1 sd         0.0900
```

d) min and max of vals above

```
country_data_2010_1980 %>%
  summarise_at(vars(rgdp_per_worker_year_1980, rgdp_per_worker_year_2010, avg_growth_labor, avg_gdp_per_worker_year_2010),
    pivot_longer(cols=everything(), names_to = "variable", values_to = "value"))
```

```
## # A tibble: 8 x 2
##   variable      value
##   <chr>      <dbl>
## 1 rgdp_per_worker_year_1980_min    967.
## 2 rgdp_per_worker_year_2010_min    968.
## 3 avg_growth_labor_min          -0.0139
## 4 avg_gdp_per_worker_min          -0.0372
## 5 rgdp_per_worker_year_1980_max 312073.
## 6 rgdp_per_worker_year_2010_max 155184.
## 7 avg_growth_labor_max           0.0899
## 8 avg_gdp_per_worker_max           0.0823
```

Savings rate (s):

```
country_data %>%
  summarise_at(vars(SavingsRate), list(min = min, max = max)) %>%
  pivot_longer(cols=everything(), names_to = "variable", values_to = "value")
```

```
## # A tibble: 2 x 2
##   variable    value
##   <chr>      <dbl>
## 1 min        0.00644
## 2 max        0.994
```