Exam2Gov355

Claire Sheppard

6/26/2020

#Clear environment  
rm(list=ls())

#Load the inequality dataset into R  
inequality\_data <- readxl::read\_xlsx("inequality.xlsx")

#view its class  
class(inequality\_data)

## [1] "tbl\_df" "tbl" "data.frame"

#view data within dataframe  
#Because the data reflects the same time period, this would be a cross sectional dataset  
head(inequality\_data)

## # A tibble: 6 x 4  
## iso2c country inequality\_gini year  
## <chr> <chr> <dbl> <dbl>  
## 1 AL Albania 32.9 2015  
## 2 AM Armenia 32.4 2015  
## 3 AT Austria 30.5 2015  
## 4 BY Belarús 25.6 2015  
## 5 BE Belgium 27.7 2015  
## 6 BZ Belize NA 2015

#view structure just to double check  
str(inequality\_data)

## tibble [203 x 4] (S3: tbl\_df/tbl/data.frame)  
## $ iso2c : chr [1:203] "AL" "AM" "AT" "BY" ...  
## $ country : chr [1:203] "Albania" "Armenia" "Austria" "Belarús" ...  
## $ inequality\_gini: num [1:203] 32.9 32.4 30.5 25.6 27.7 NA 47.8 NA NA 46.7 ...  
## $ year : num [1:203] 2015 2015 2015 2015 2015 ...

# The data frame inequality\_gini Corresponds to the inequality Gini index, which “measures the extent to which the distribution of income (or, in some cases, consumption expenditure) among individuals or households within an economy deviates from a perfectly equal distribution.   
inequality\_data[complete.cases(inequality\_data),]

## # A tibble: 80 x 4  
## iso2c country inequality\_gini year  
## <chr> <chr> <dbl> <dbl>  
## 1 AL Albania 32.9 2015  
## 2 AM Armenia 32.4 2015  
## 3 AT Austria 30.5 2015  
## 4 BY Belarús 25.6 2015  
## 5 BE Belgium 27.7 2015  
## 6 BJ Benin 47.8 2015  
## 7 BO Bolivia 46.7 2015  
## 8 BW Botswana 53.3 2015  
## 9 BR Brazil 51.9 2015  
## 10 BG Bulgaria 38.6 2015  
## # ... with 70 more rows

inequality\_data

## # A tibble: 203 x 4  
## iso2c country inequality\_gini year  
## <chr> <chr> <dbl> <dbl>  
## 1 AL Albania 32.9 2015  
## 2 AM Armenia 32.4 2015  
## 3 AT Austria 30.5 2015  
## 4 BY Belarús 25.6 2015  
## 5 BE Belgium 27.7 2015  
## 6 BZ Belize NA 2015  
## 7 BJ Benin 47.8 2015  
## 8 BM Bermuda NA 2015  
## 9 BT Bhutan NA 2015  
## 10 BO Bolivia 46.7 2015  
## # ... with 193 more rows

#It is better to have a lower inequality\_gini score

head(inequality\_data)

## # A tibble: 6 x 4  
## iso2c country inequality\_gini year  
## <chr> <chr> <dbl> <dbl>  
## 1 AL Albania 32.9 2015  
## 2 AM Armenia 32.4 2015  
## 3 AT Austria 30.5 2015  
## 4 BY Belarús 25.6 2015  
## 5 BE Belgium 27.7 2015  
## 6 BZ Belize NA 2015

#remove accents  
remove.accents <- function(s) {  
# 1 character substitutions  
old1 <- "ús"  
new1 <- "us"  
s1 <- chartr(old1, new1, s)  
}  
#finish removal  
inequality\_data$country = remove.accents(inequality\_data$country)

decreasing<- sort(inequality\_data$inequality\_gini,decreasing = FALSE)  
head(decreasing,5)

## [1] 25.4 25.5 25.6 25.9 26.5

inequality\_data2 <- inequality\_data[complete.cases(inequality\_data),]  
mean = mean(inequality\_data2$inequality\_gini)  
mean

## [1] 36.81375

#I truly am an intelligent being. I went through 6-7-8 AND 3 in the past few days. Lost 8 hours after my laptop crashed trying to reinstall everything and recreate notes. Hoping this works at least for partial credit again. Sadly I read the questions and thought, Yeah I know this :(

Note that the echo = FALSE parameter was added to the code chunk to prevent printing of the R code that generated the plot.