Exam #2

Braden Holt

6/26/2020

## Questions

1. Clear Environment

rm(list=ls(all = TRUE))

1. Load Data

inequality\_data <- import("inequality.xlsx")

1. This is a cross sectional dataset because it captures a single observation for each variable.

cross.sectional <- inequality\_data$year != 2015  
print(cross.sectional)

## [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [25] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [37] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [49] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [61] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [73] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [85] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [97] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [109] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [121] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [133] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [145] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [157] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [169] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [181] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [193] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

1. It is better to have a lower inequality score because the more unequal country of the three, Brazil, has a higher score.

inequality\_gini\_denmark <- subset(inequality\_data, country == 'Denmark', select = inequality\_gini)  
inequality\_gini\_sweden <- subset(inequality\_data, country == 'Sweden', select = inequality\_gini)  
inequality\_gini\_brazil <- subset(inequality\_data, country == 'Brazil', select = inequality\_gini)

1. Mean inequality\_gini score is 36.81375

mean(na.omit(inequality\_data$inequality\_gini))

## [1] 36.81375

1. If Else Commands and Dummy Variables

inequality\_data$high\_inequality = NA  
inequality\_data$low\_inequality = NA  
  
inequality\_data$high\_inequality <-  
 if(inequality\_data$inequality\_gini > 36.81375){  
 inequaity\_data$high\_inequality = 1  
 } else if (inequality\_data$inequality\_gini < 36.81375) {  
 inequality\_data$high\_inequality = 0  
 }  
  
inequality\_data$low\_inequality <-   
 if(inequality\_data$inequality\_gini < 36.81375){  
 inequality\_data$low\_inequality = 1  
 } else if (inequality\_data$inequality\_gini < 36.81375) {  
 inequality\_data$low\_inequality = 0  
 }

1. For Loop

x <- c("The World Bank", "African Development Bank", "Bill and Melinda Gates Foundation")  
  
for(i in x){  
 print(x)  
}

## [1] "The World Bank" "African Development Bank"   
## [3] "Bill and Melinda Gates Foundation"  
## [1] "The World Bank" "African Development Bank"   
## [3] "Bill and Melinda Gates Foundation"  
## [1] "The World Bank" "African Development Bank"   
## [3] "Bill and Melinda Gates Foundation"

1. Import WDI Data

gdp\_per\_cap <- WDI(country = 'all',  
 indicator = c('NY.GDP.PCAP.CD'),  
 start = 2015,  
 end = 2015,   
 cache = NULL,  
 extra = FALSE)

1. Rename Variables

setnames(gdp\_per\_cap, "NY.GDP.PCAP.CD", "GDP Per Capita")

1. Merge Data

## Joining, by = c("iso2c", "country", "year")

1. Remove Missing Data

merged\_df<- merged\_df %>%   
 dplyr::filter(!(inequality\_gini == "NA"))

1. Keep Inequality Scores > 30

data\_greater\_30 <-   
 merged\_df %>%   
 dplyr::filter(inequality\_gini > 30)

1. Sum of ‘ai’

sum(grep("ai", data\_greater\_30))

## [1] 2

1. Apply Function

sapply(data\_greater\_30$inequality\_gini, sum)

## [1] 32.9 32.4 30.5 47.8 46.7 53.3 51.9 38.6 42.4 44.4 38.6 51.1 48.4 41.5 31.1  
## [16] 34.0 45.2 46.0 31.8 40.6 32.7 35.0 32.7 35.9 36.5 31.7 36.0 49.6 30.4 41.0  
## [31] 39.5 31.8 35.4 40.8 34.2 37.4 33.8 41.0 39.0 38.1 59.1 35.6 33.5 50.8 47.6  
## [46] 43.4 44.4 31.8 35.5 35.9 37.7 40.5 36.2 32.3 34.0 36.0 43.1 37.6 32.8 42.9  
## [61] 33.2 40.1 57.1

1. Label

var\_label(merged\_df) <- list('country' = "Country",  
 'iso2c' = "ISO2C Country Code",  
 'inequality\_gini' = "Inequality Score",  
 'year' = "Year",  
 'high\_inequality' = "High Inequality",  
 'low\_inequality' = "Low Inequality",  
 'GDP Per Capita' = "GDP Per Capita")