Week 5 In-class Assignment

Felix Ho

2024-10-07

Create Table 1 for armed conflict paper.

finaldata <- read.csv(here("data", "mergealldata.csv"), header = TRUE)   
baseline <- finaldata %>%   
 filter(year == 2000)  
baseline$OECDf <- factor(baseline$OECD, levels = c(1,0),   
 labels = c("Yes", "No"))  
baseline$droughtf <- factor(baseline$drought, levels = c(1,0),   
 labels = c("Presence", "Absence"))  
baseline$earthquakef <- factor(baseline$earthquake, levels = c(1,0),   
 labels = c("Presence", "Absence"))  
baseline$armconf <- factor(baseline$armcon, levels = c(1,0),   
 labels = c("Yes", "No"))  
  
label(baseline$gdp1000) <- "GDP per capita"  
label(baseline$OECDf) <- "OECD member"  
label(baseline$popdens) <- "Population density"  
label(baseline$urban) <- "Urban residence"  
label(baseline$agedep) <- "Age dependency ratio"  
label(baseline$male\_edu) <- "Male education"  
label(baseline$temp) <- "Mean population-weighted annual temperature"  
label(baseline$rainfall1000) <- "Mean population-weighted annual rainfall"  
label(baseline$droughtf) <- "Droughts"  
label(baseline$earthquakef) <- "Earthquakes"  
  
units(baseline$gdp1000) <- "USD"  
  
# Set up the rows or labels of the table. Group the two "Yes" and "No" strata  
# under a common heading" "Armed conflict exposure".  
  
labels <- list(  
 variables = list(gdp1000 = render.varlabel(baseline$gdp1000),  
 OECDf = render.varlabel(baseline$OECDf),  
 popdens = render.varlabel(baseline$popdens),  
 urban = render.varlabel(baseline$urban),  
 agedep = render.varlabel(baseline$agedep),  
 male\_edu = render.varlabel(baseline$male\_edu),  
 temp = render.varlabel(baseline$temp),  
 rainfall1000 = render.varlabel(baseline$rainfall1000),  
 droughtf = render.varlabel(baseline$droughtf),  
 earthquakef = render.varlabel(baseline$earthquakef)  
 ),   
 groups=list("", "Armed conflict in 2000"))  
  
# Set up the strata or columns of the table.  
  
strata <- c(list(Total=baseline), split(baseline, baseline$armconf))  
  
# Make all values in the table have 1 decimal place.  
  
my\_summary <- function(x) {  
 # Remove NAs before calculation.  
 x <- na.omit(x)  
 # Calculate median, 25th and 75th percentiles.  
 q <- quantile(x, probs = c(0.25, 0.5, 0.75))  
 # Format the output as "median [25th%, 75th%]" with 1 decimal place.  
 sprintf("%.1f [%.1f, %.1f]", q[2], q[1], q[3])  
}  
  
table1(strata, labels, groupspan=c(1, 2),   
 caption = "Table 1: Description of data used in the study",   
 footnote = "Data given as median [25th percentile, 75th percentile] for continuous variables or counts (%) for categorical variables.",  
 render.continuous = my\_summary)

|  |  | Armed conflict in 2000 | |
| --- | --- | --- | --- |
|  | Total (N=186) | Yes (N=39) | No (N=147) |
| **GDP per capita (USD)** | 1.8 [0.6, 6.1] | 0.6 [0.3, 1.1] | 2.2 [0.9, 11.0] |
| Missing | 5 (2.7%) | 2 (5.1%) | 3 (2.0%) |
| **OECD member** |  |  |  |
| Yes | 30 (16.1%) | 1 (2.6%) | 29 (19.7%) |
| No | 156 (83.9%) | 38 (97.4%) | 118 (80.3%) |
| **Population density** | 25.4 [13.4, 37.7] | 21.3 [11.7, 32.0] | 27.3 [13.6, 39.4] |
| Missing | 1 (0.5%) | 0 (0%) | 1 (0.7%) |
| **Urban residence** | 28.0 [15.4, 39.7] | 24.1 [19.3, 32.2] | 28.9 [14.8, 41.4] |
| Missing | 1 (0.5%) | 0 (0%) | 1 (0.7%) |
| **Age dependency ratio** | 63.5 [49.5, 84.7] | 84.4 [62.3, 93.6] | 60.2 [48.5, 78.5] |
| **Male education** | 7.1 [4.9, 10.1] | 4.9 [3.1, 6.8] | 7.9 [5.6, 10.6] |
| Missing | 1 (0.5%) | 0 (0%) | 1 (0.7%) |
| **Mean population-weighted annual temperature** | 21.4 [13.0, 25.4] | 24.0 [19.3, 26.5] | 21.0 [11.3, 25.2] |
| Missing | 1 (0.5%) | 0 (0%) | 1 (0.7%) |
| **Mean population-weighted annual rainfall** | 1.0 [0.6, 1.6] | 1.1 [0.5, 1.6] | 1.0 [0.6, 1.7] |
| Missing | 1 (0.5%) | 0 (0%) | 1 (0.7%) |
| **Droughts** |  |  |  |
| Presence | 22 (11.8%) | 3 (7.7%) | 19 (12.9%) |
| Absence | 164 (88.2%) | 36 (92.3%) | 128 (87.1%) |
| **Earthquakes** |  |  |  |
| Presence | 18 (9.7%) | 5 (12.8%) | 13 (8.8%) |
| Absence | 168 (90.3%) | 34 (87.2%) | 134 (91.2%) |
| Data given as median [25th percentile, 75th percentile] for continuous variables or counts (%) for categorical variables. | | | |

Create a figure that shows the trend in maternal mortality for countries that had an increase from 2000 to 2017. First, create a new variable diffmatmor that shows the difference between maternal mortality in 2017 and maternal mortality in 2000.

graphdata <- finaldata %>%   
 # Only select the variables needed to create the figure.  
 select(country\_name, year, matmor) %>%   
 # Select years 2000 to 2017.  
 filter(year == c(2000, 2017)) %>%   
 # Sort data by country and increasing year.  
 arrange(country\_name, year) %>%   
 # Group by country.  
 group\_by(country\_name) %>%   
 # Create a new variable diffmatmor that takes the difference between each year and the baseline year.  
 mutate(diffmatmor = matmor - matmor[1L]) %>%  
 # Only include countries that had an increase in mortality from 2000 to 2017.  
 filter(diffmatmor > 0)  
  
finaldatag <- finaldata %>%   
 filter(country\_name %in% graphdata$country\_name)  
  
finaldatag %>%  
 ggplot(aes(x = year, y = matmor, group = country\_name)) +  
 geom\_line(aes(color = as.factor(country\_name)), alpha = 0.5, lwd = 0.7) +  
 xlim(c(2000,2017)) +  
 # Use log 10 scale for y axis.  
 scale\_y\_continuous(trans='log10') +  
 labs(y = "Maternal mortality (log 10 scale)", x = "Year", color = "Countries", title = "Maternal mortality trend of countries with increase from 2000-2017")

