Question 1

2025-11-16

Data

Just a brief description of the data used in this analysis. There are 4 files which I am going to import to this file later, two for Kenya, two for Bangladesh.

For Kenya, there are "Kenya_Wealth", "Kenya_Income", and in each of them, it contains the following indicators:

- Gini coefficient
- Share of income/wealth held by highest 1%
- Share of income/wealth held by highest 10%
- Share of income/wealth held by bottom 50%
- Share of income/wealth held by middle 40%

The same applies for Bangladesh, with files "Bangladesh_Wealth" and "Bangladesh_Income".

Note that the data for Bangladesh are imported in form of .xlsx files, as I found that it would be easier to import than the .csv file generated by the website as csv files generated are not in regular format. And I have changed the name for the files for convenience and clarity of the files.

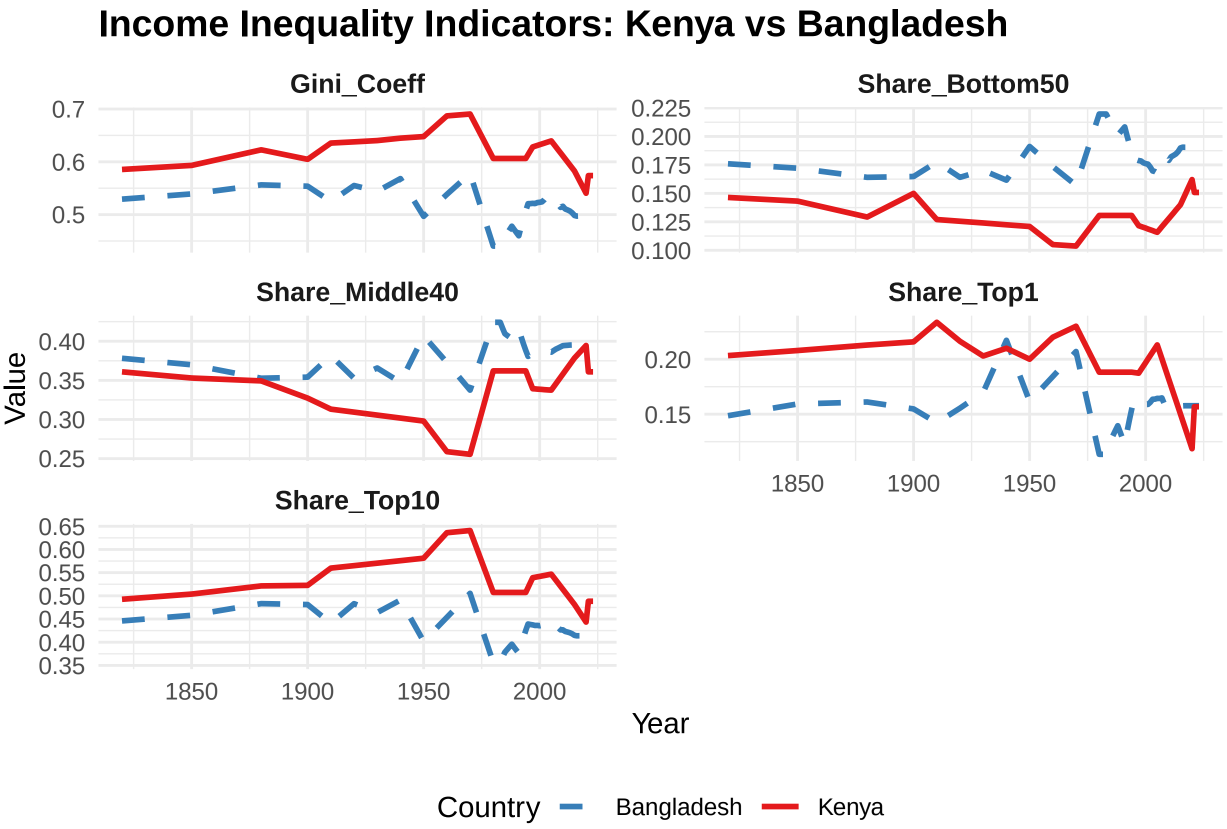
```r
# !!! DON"T FORGET TO CHANGE THE WORKING DIRECTORY TO YOUR OWN DIRECTORY !!!
setwd("~/R/BI_Group_Proj/Data")

# Importing Kenya Data
Ken_W_Ineq <- read_delim("Kenya_Wealth_Inequality.csv",
  delim = ";", escape_double = FALSE, col_names = FALSE,
  trim_ws = TRUE, skip = 1
) # This is why I would rather use xls file

Ken_I_Ineq <- read_delim("Kenya_Income_Inequality.csv",
  delim = ";", escape_double = FALSE, col_names = FALSE,
  trim_ws = TRUE, skip = 1
)

# Importing Bangladesh Data
Bang_W_Ineq <- read_excel("Bang_Wealth_Inequality.xlsx", col_names = FALSE)
Bang_I_Ineq <- read_excel("Bang_Income_Inequality.xlsx", col_names = FALSE)
```

```r
# Make a function for cleaning data sets
clean_data_inequality <- function(x) {
  colnames(x) <- c("Country", "Indicator", "Percentile", "Year", "Value")
  # Columns of Inequality data sets are all in this order, check when use for others
  x <- x %>%
    pivot_wider(
      names_from = Percentile,
      values_from = Value
    ) %>%
    filter(!if_all(c(pall, p0p50, p50p90, p90p100, p99p100), is.na)) %>%
    # filter out those rows where all the values are NA
    select(Country, Year, pall, p0p50, p50p90, p90p100, p99p100) %>%
    # To ensure the columns are in correct order and delete indicator column
    group_by(Country, Year) %>%
    summarise(across(
      c(pall, p0p50, p50p90, p90p100, p99p100),
      ~ first(na.omit(.))
    ),
    .groups = "drop")
    # Note that previously we have 5 lines for a single year, and each
    # line only shows a single indicator. By doing this, we combine the data together.
  colnames(x) <- c(
    "Country", "Year", "Gini_Coeff", "Share_Bottom50",
    "Share_Middle40", "Share_Top10", "Share_Top1"
  )
  return(x)
}

Bang_I_Ineq_wider <- clean_data_inequality(Bang_I_Ineq)
Bang_W_Ineq_wider <- clean_data_inequality(Bang_W_Ineq)
Ken_I_Ineq_wider <- clean_data_inequality(Ken_I_Ineq)
Ken_W_Ineq_wider <- clean_data_inequality(Ken_W_Ineq)

longer_format <- function(x) {
  x %>% pivot_longer(
    cols = c(
      Gini_Coeff, Share_Bottom50, Share_Middle40,
      Share_Top10, Share_Top1
    ),
    names_to = "Indicator",
    values_to = "Value"
  )
}

Bang_I_Ineq_longer <- longer_format(Bang_I_Ineq_wider)
Bang_W_Ineq_longer <- longer_format(Bang_W_Ineq_wider)
Ken_I_Ineq_longer <- longer_format(Ken_I_Ineq_wider)
Ken_W_Ineq_longer <- longer_format(Ken_W_Ineq_wider)
```

Comparing Kenya with Bangladesh

In this sector I will draw some line chart comparing inequality of income & wealth between Kenya and Bangladesh by means of Gini coefficient and share of income/wealth occupied by different social class.

```
# Gini Coefficient
I_Inequality_K_and_B <- bind_rows(Ken_I_Ineq_longer, Bang_I_Ineq_longer)

I_Inequality_K_and_B %>% ggplot(aes(x = Year, y = Value, color = Country, linetype = Country)) +
  geom_line(linewidth = 1) +
  facet_wrap(~Indicator, scales = "free_y", ncol = 2) +
  scale_color_manual(values = c("Kenya" = "#E41A1C", "Bangladesh" = "#377EB8")) +
  scale_linetype_manual(values = c("Kenya" = "solid", "Bangladesh" = "dashed")) +
  labs(
    title = "Income Inequality Indicators: Kenya vs Bangladesh",
    x = "Year",
    y = "Value"
  ) +
  theme_minimal() +
  theme(
    legend.position = "bottom",
    plot.title = element_text(face = "bold", size = 14),
    strip.text = element_text(face = "bold", size = 10)
  )
```


Income Inequality Indicators: Kenya vs Bangladesh
Gini_Coeff
Share_Bottom50
Share_Middle40
Share_Top1
Share_Top10
Value
Year
Country
Bangladesh
Kenya