WP4 Submission

Meghan Edgerton

2023-02-10

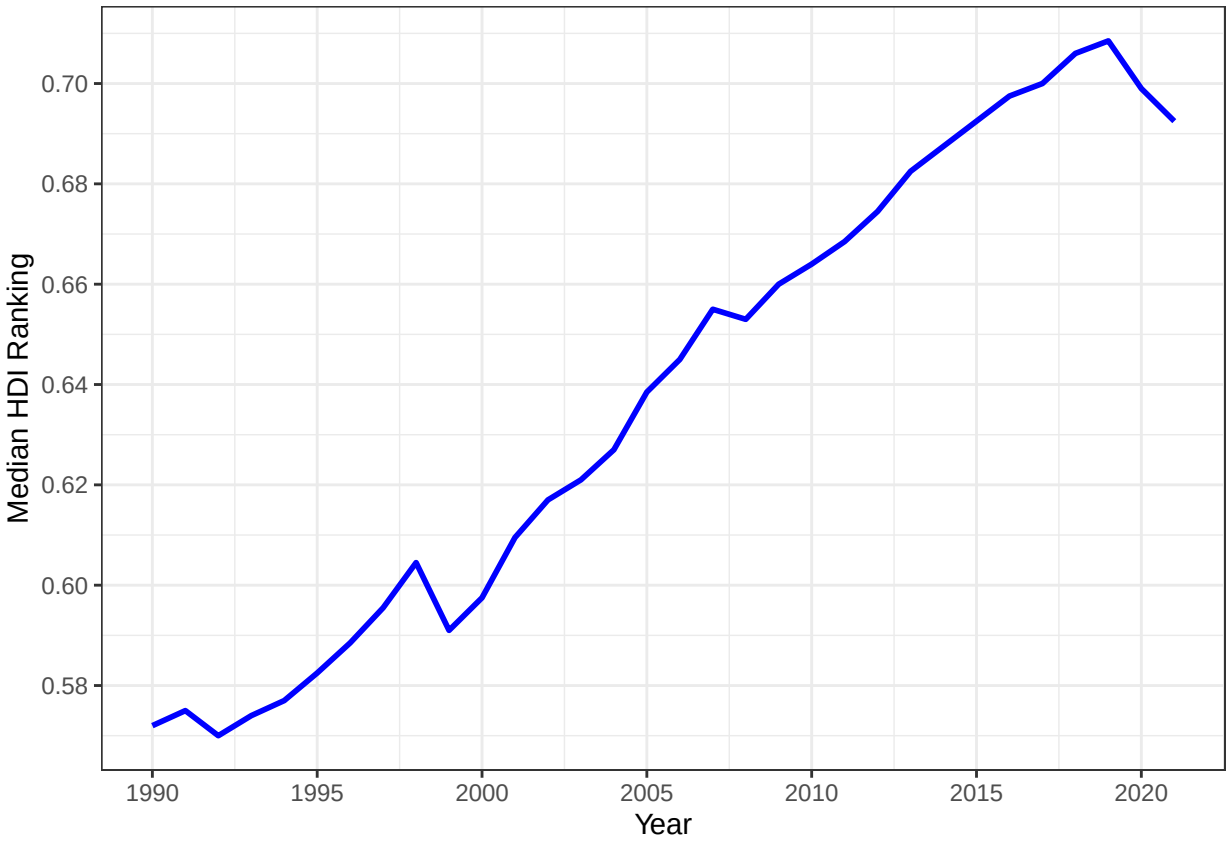

Figure 1. Line plot displaying the general increase in the annual median Human Development Index (HDI) ranking in different regions of the world from the years of 1990 to 2021. This data was pulled from the United Nation's Human Development Reports website, where it was filtered to show annual HDI rankings for countries assigned to one of the following six regions: Arab States (AS), East Asia and the Pacific (EAP), Europe and Central Asia (ECA), Latin America and the Caribbean (LAC), South Asia (SA), and Sub-Saharan Africa (SSA). The median HDI ranking was then calculated for each year from 1990 to 2021.

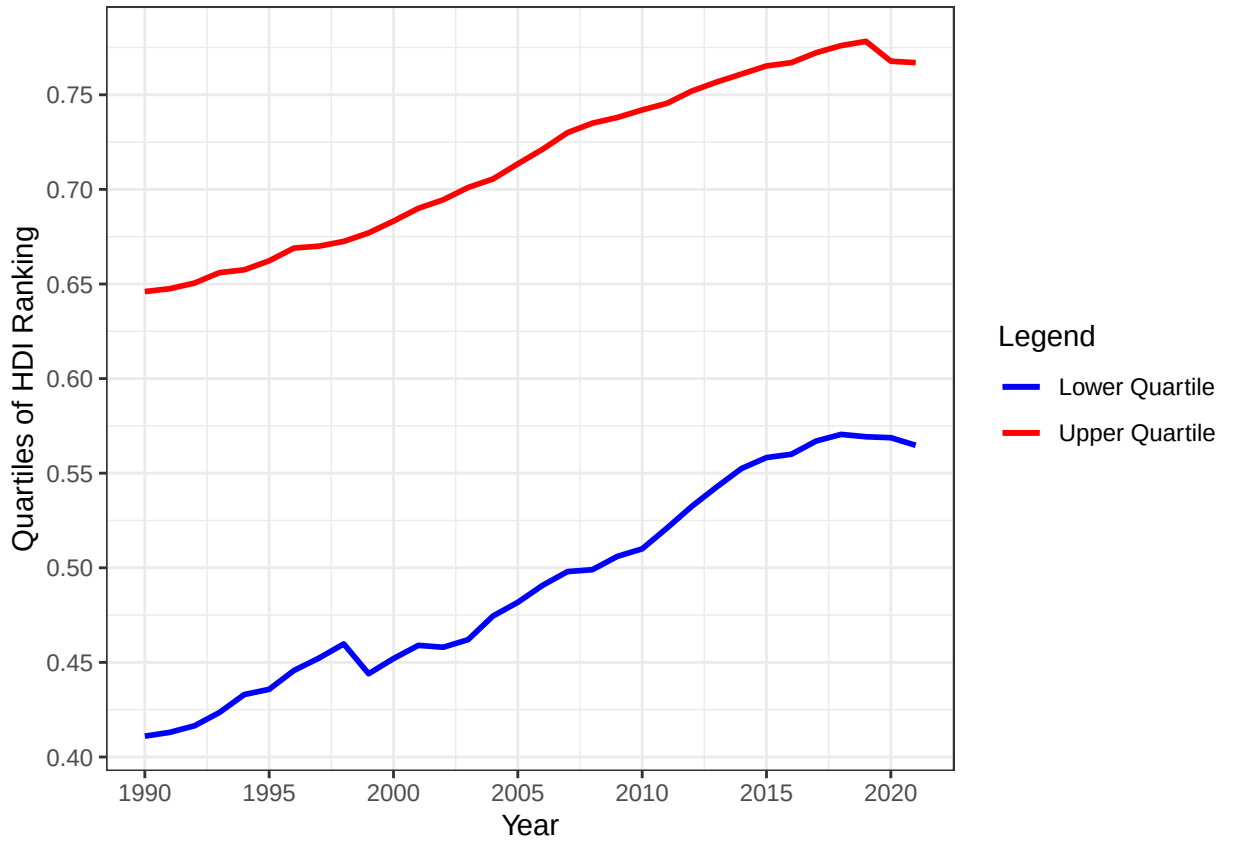

Figure 2. Line plot displaying the general increase in the lower and upper quartiles of the annual Human Development Index (HDI) ranking in different regions of the world from the years of 1990 to 2021. This data was pulled from the United Nation's Human Development Reports website, where it was filtered to show annual HDI rankings for countries assigned to one of the following six regions: Arab States (AS), East Asia and the Pacific (EAP), Europe and Central Asia (ECA), Latin America and the Caribbean (LAC), South Asia (SA), and Sub-Saharan Africa (SSA). The lower and upper quartiles of the annual HDI rankings were then calculated for each year from 1990 to 2021.

R Code Used To Make These Plots

```
library(dplyr)
library(ggplot2)
library(rmarkdown)
library(gridExtra)
library(scales)
library(data.table)

path <- "/Users/meghanedgerton/Documents/DataVizINFO526/HDI Edited.csv"

data <- read.csv(path)
```

```
dt <- data.table(Country = data$country,
                 Region = data$region,
                 Year = data$year,
                 HDI = data$annual_hdi)

# Filtering out rows with NULL Region values
yearlydata <- dt %>% filter(Region != "")  %>% group_by(Year) %>%
  summarize(avg = mean(HDI),
            stdev = sd(HDI),
            med = median(HDI),
            lowerQ = quantile(HDI, probs = 0.25),
            upperQ = quantile(HDI, probs = 0.75))

cols <- c("Lower Quartile" = "blue", "Upper Quartile" = "red" )
```

```
# Central Tendency Plot

p1 <- ggplot(yearlydata, aes(x = Year, y = med)) +
  geom_line(linewidth = 1, color = "blue") +
  theme_bw() +
  scale_x_continuous(limits = c(1990,2021), n.breaks = 6) +
  scale_y_continuous(n.breaks = 8) +
  labs(x = "Year", y = "Median HDI Ranking")
```

```
# Variation Plot
p2 <- ggplot() +
  geom_line(data = yearlydata, aes(x=Year, y=lowerQ, color = "Lower Quartile"),
            linewidth = 1) +
  geom_line(data = yearlydata, aes(x=Year, y=upperQ, color = "Upper Quartile"),
            linewidth = 1) +
  theme_bw() +
  scale_x_continuous(limits = c(1990,2021), n.breaks = 6) +
  scale_y_continuous(n.breaks = 8) +
  labs(x = "Year", y = "Quartiles of HDI Ranking", color = "Legend") +
  scale_color_manual(values = cols)
```