Exploring Infant Mortality

Nourhan Ghanima

2023-11-27

Answering Some Questions Related to Infant Mortality Rates Around the World

Infant mortality refers to the death of infants, typically those under the age of one year. It is a measure of the number of deaths that occur in infants within a given population or region. Infant mortality rate is commonly expressed as the number of deaths per 1,000 live births. It is an important health indicator, reflecting the overall health and well-being of a population, access to healthcare services, socio-economic conditions, and other factors. Reducing infant mortality is a goal of public health efforts worldwide.

In this very brief project I hope to answer some questions about the global rates of infant deaths, namely:

- what factors affect infant mortality and in which direction.

- are there differences in infant mortality at the continent level?

- which regions have the highest infant mortality around the world.

I start by loading the data, sourced from UNICEF, and then by tidying up the data by ensuring:

- that each variable has its own column

- that each observation has its own row

- that each value has its own cell

```r
library(readr)
library(tidyverse)
library(dplyr)
library(ggplot2)
library(readxl)
```

```r
mortality_data <- read_csv("/Users/nourhanghanima/downloads/T1.csv", skip = 1,
                           show_col_types = FALSE)
```

```r
mortality_data <- mortality_data |>
                  rename(Country = "Country Name") |>
                  pivot_longer("1960":"2021", names_to = "year",
                               values_to = "mortality_rate")

mortality_data$year <- as.numeric(mortality_data$year)

summary(mortality_data)
```

```
##    Country          Country Code       Indicator Name      Indicator Code
##  Length:7254        Length:7254        Length:7254         Length:7254
##  Class :character   Class :character   Class :character    Class :character
##  Mode  :character   Mode  :character   Mode  :character    Mode  :character
##
##
```

```
##
##       year        mortality_rate
##  Min.   :1960   Min.   :  1.70
##  1st Qu.:1975   1st Qu.: 13.40
##  Median :1990   Median : 31.80
##  Mean   :1990   Mean   : 46.91
##  3rd Qu.:2006   3rd Qu.: 68.30
##  Max.   :2021   Max.   :220.40
```

I then create a new column for continent, in order to perform analysis at the continent level.

```
mortality_data <- mortality_data |>
  mutate(Continent = case_when(
    Country %in% c("Algeria", "Angola",  "Benin", "Botswana", "Burkina Faso",
                   "Burundi","Cabo Verde/Cape Verde","Cameroon", "Central African Republic",
                   "Chad","Comoros", "Congo", "Democratic Republic of the Congo", "Djibouti",
                   "Egypt", "Equatorial Guinea","Eritrea","Eswatini",
                   "Ethiopia","Gabon","Gambia","Ghana", "Guinea", "Guinea-Bissau",
                   "Ivory Coast", "Kenya", "Lesotho", "Liberia",
                   "Libya", "Madagascar", "Malawi", "Mali", "Mauritania", "Mauritius",
                   "Morocco", "Mozambique", "Namibia", "Niger", "Nigeria", "Rwanda",
                   "Brunei Darussalam", "Sao Tome and Principe","Senegal", "Seychelles",
                   "Sierra Leone", "Somalia", "South Africa", "South Sudan", "Sudan",
                   "Tanzania", "Togo", "Tunisia","Uganda", "Zambia", "Zimbabwe",
                   "Cote d'Ivoire", "Congo, Rep.",
                   "Cabo Verde", "Egypt, Arab Rep.", "Gambia, The") ~ "Africa",
    Country %in% c("Afghanistan", "Armenia", "Azerbaijan", "Bahrain", "Bangladesh",
                   "Bhutan", "Brunei", "Cambodia", "China", "East Timor", "Georgia",
                   "India", "Indonesia", "Iran", "Iraq", "Israel", "Japan", "Jordan",
                   "Kazakhstan", "Kuwait", "Kyrgygstan", "Laos", "Lebanon", "Malaysia",
                   "Maldives", "Mongolia", "Myanmar", "Nepal", "North Korea", "Oman",
                   "Pakistan","Palestine", "Philippines", "Qatar", "Russia", "Vietnam",
                   "Saudi Arabia", "Singapore","South Korea", "Sri Lanka", "Syria",
                   "Syrian Arab Republic", "Taiwan", "Tajikistan", "Thailand", "Turkey",
                   "Turkiye","Turkmenistan", "United Arab Emirates", "Uzbekistan", "Yemen",
                   "Korea, Rep.") ~ "Asia",
    Country %in% c("Albania", "Andorra", "Austria", "Belarus","Belgium",
                   "Bosnia and Herzegovina", "Bulgaria", "Croatia", "Cyprus",
                   "Czech Republic","Denmark", "Estonia", "Finland", "France",
                   "Georgia", "Germany","Greece", "Hungary", "Iceland", "Ireland",
                   "Italy", "Kosovo", "Latvia", "Liechtenstein","Lithuania",
                   "Luxembourg", "North Macedonia", "Malta", "Moldova", "Monaco",
                   "Montenegro", "Netherlands", "Norway", "Poland", "Portugal",
                   "Romania", "Russia", "San Marino", "Serbia", "Slovakia",
                   "Slovenia", "Spain","Sweden", "Switzerland", "Turkey", "Ukraine",
                   "United Kingdom", "Vatican City") ~ "Europe",
    Country %in% c("Antigua and Barbuda", "Bahamas", "Barbados", "Belize", "Canada",
                   "Costa Rica", "Cuba", "Dominica", "Dominican Republic", "El Salvador",
                   "Grenada", "Guatemala", "Haiti", "Honduras", "Jamaica", "Mexico",
                   "Nicaragua", "Panama", "Saint Kitts and Necis","Saint Lucia",
                   "Saint Vincent and the Grenadines", "Trinidad and Tobago", "United States",
                   "St. Lucia", "Turks and Caicos Islands", "St. Vincent and the Grenadines",
                   "British Virgin Islands") ~ "North America",
    Country %in% c("Argentina", "Bolivia","Brazil", "Chile", "Colombia", "Ecuador",
                   "Guyana", "Paraguay", "Peru", "Suriname", "Trinidad and Tobago",
```

```
                "Uruguay", "Venezuela", "Venezuela, RB" ) ~ "South America",
  Country %in% c("Australia", "Fiji", "Kiribati", "Marshall Islands",
                "Federated States of Micronesia", "Nauru", "New Zealand",
                "Palau",  "Papua New Guinea", "Samoa", "Solomon Islands", "Tonga",
                "Tuvalu", "Vanuatu", "Micronesia, Fed. Sts.") ~ "Oceania"))
```

Plots

The first plot in this report shows the mortality rate for each country across different years, color coded by each continent. It is obvious that while mortality rates are decreasing everywhere, infant mortality is still significantly higher in Africa than in other continents.
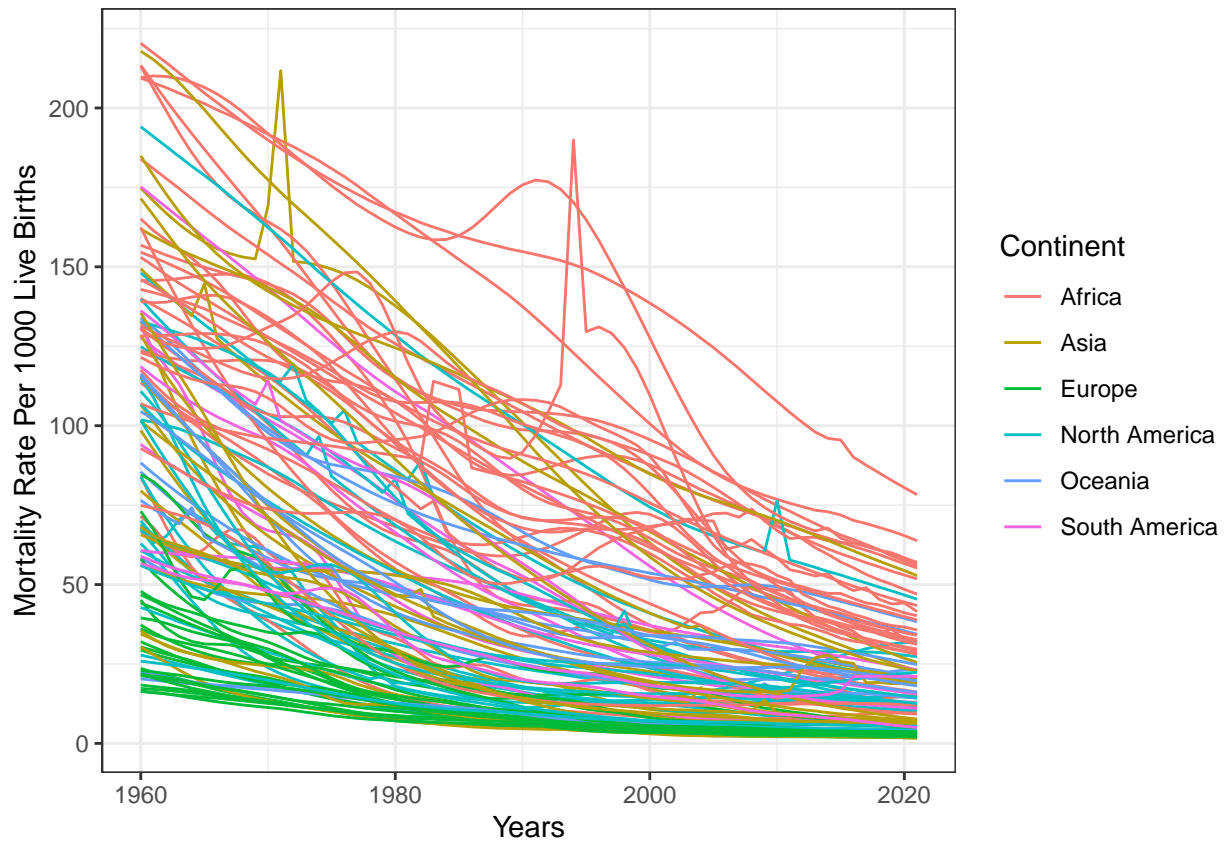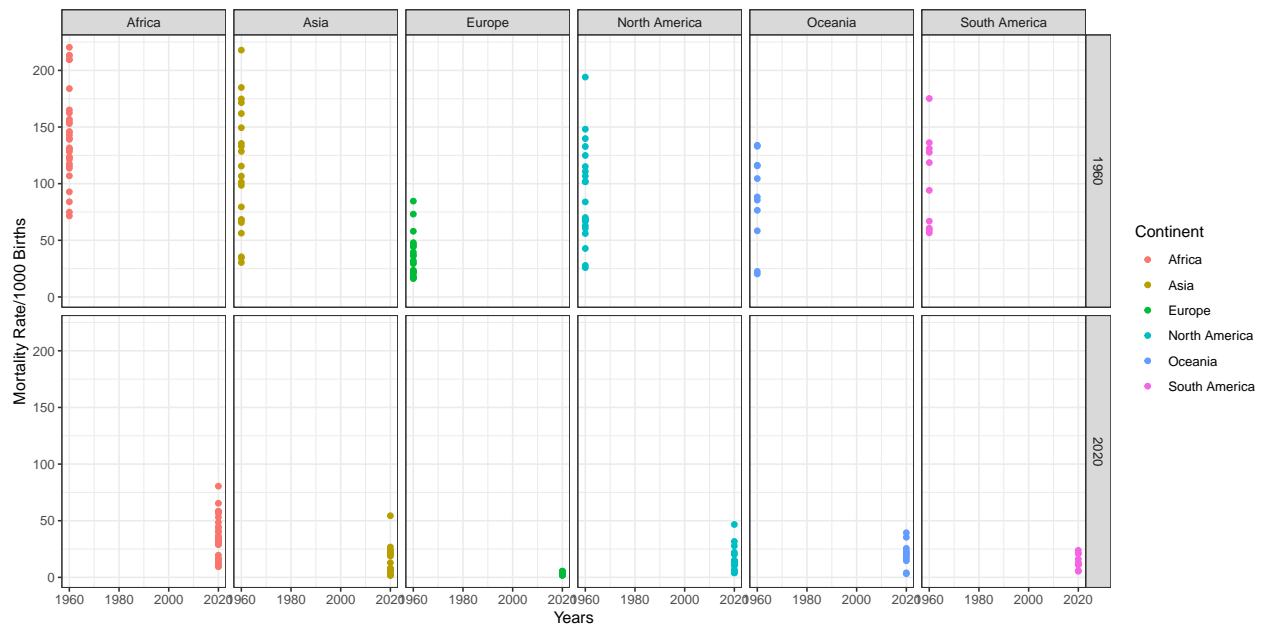


```
## <ScaleContinuousPosition>
##  Range:
##  Limits: 1.96e+03 -- 2.02e+03
```

The second plot shows the mortality rates for each continent in the years 1960 and 2020, which show the first year in the data and the last year in the data, respectively.

The figures again show a marked decrease in infant mortality between 1960 and 2020, but the rates for Africa are still significantly higher than the other continents.

Which country has the highest infant mortality for the year 2020?

```
## # A tibble: 1 x 4
##   Country        year mortality_rate Continent
##   <chr>         <dbl>          <dbl> <chr>
## 1 Sierra Leone   2020           80.5 Africa
```

Which countries have the lowest infant mortality for the year 2020?

```
## # A tibble: 4 x 4
##   Country     year mortality_rate Continent
##   <chr>      <dbl>          <dbl> <chr>
## 1 Finland     2020            1.8 Europe
## 2 Japan       2020            1.8 Asia
## 3 Norway      2020            1.8 Europe
## 4 Singapore   2020            1.8 Asia
```

How is the progression of infant mortality different for the highest and lowest countries over the years?
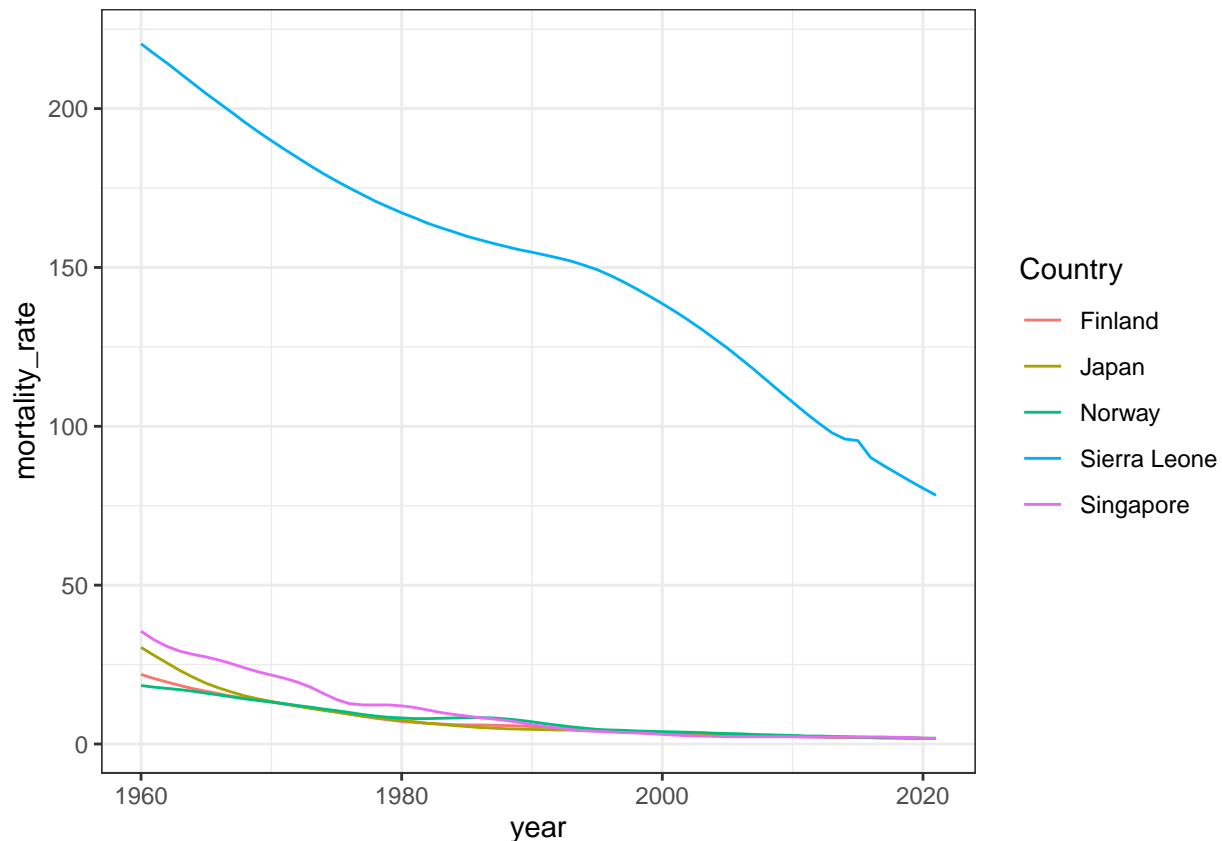
Infant mortality is decreasing everywhere along the years, but for Sierra Leone even the lowest infant mortality is higher than the highest infant mortality in Japan, Norway, Finland, or Singapore.

Can we figure out why infant mortality in Sierra Leone is so high?

Before looking at Sierra Leone specifically, I'd like to look into the factors that affect infant mortality and the directions of the relationships more generally.

To do this, I'll add data to the original mortality dataset that includes the rates of HIV prevalence, GDP, public health expenditure, and fertility rates.

After combining the data, I run a simple regression model to check whether HIV rates, fertility rates, public spending, and GDP have any effects on infant mortality.

```
mortality_model <- lm(mortality_rate ~ hiv_rate + fertility_rate + GDP + public_spending, data= df2)
summary(mortality_model)
```

```
##
## Call:
## lm(formula = mortality_rate ~ hiv_rate + fertility_rate + GDP +
##     public_spending, data = df2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -28.441  -4.848  -1.342   3.620  59.159
##
```

```
## Coefficients:
##                  Estimate Std. Error t value Pr(>|t|)
## (Intercept)     -8.237e+00  1.901e+00  -4.334 1.68e-05 ***
## hiv_rate         9.607e-01  8.779e-02  10.943  < 2e-16 ***
## fertility_rate   1.354e+01  3.729e-01  36.322  < 2e-16 ***
## GDP             -2.645e-14  1.912e-13  -0.138     0.89
## public_spending -2.007e-01  3.897e-02  -5.152 3.38e-07 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 9.662 on 689 degrees of freedom
## Multiple R-squared:  0.8017, Adjusted R-squared:  0.8006
## F-statistic: 696.4 on 4 and 689 DF,  p-value: < 2.2e-16
```

We can see from the results that HIV rates have a highly statistically significant positive relationship with infant mortality. The same is true for fertility rates. Public spending, on the other hand, has a highly statistically significant effect on mortality rates. As public spending increases, infant mortality decreases. These results are expected.