Understanding Life Expectancy and Health Disparities between Developed and Developing Countries

by Alissa, Cathy, Caroline, and Srishti

Topic and Motivation

Topic: Life expectancy and factors that influence it

- Predictors of life expectancy
- Developed vs Developing countries

Research Question: Which health and socioeconomic factors most greatly affect the life expectancy of a country, and does this differ between developed and developing countries?

Motivation:

- Important marker for overall health of a population
- Underlying health inequities across countries
- Predictors of life expectancy can inform crisis preparedness

Life Expectancy Dataset

**Origin of dataset:** World Health Organization (WHO)

- Each observation is a country's health status for a given year in 2000-2015

    - Includes variables such as life expectancy, mortality rates, country development status, disease rates, etc.

- Data is collected through household surveys and health facilities

- Ethical considerations of health/medical data

Highlights from EDA

```
# A tibble: 17 × 2
   name                                               value
   <chr>                                              <dbl>
 1 cor(Life.expectancy, HIV.AIDS)                    -0.592
 2 cor(Life.expectancy, thinness..1.19.years)        -0.458
 3 cor(Life.expectancy, thinness.5.9.years)          -0.458
 4 cor(Life.expectancy, under.five.deaths)           -0.192
 5 cor(Life.expectancy, infant.deaths)               -0.169
 6 cor(Life.expectancy, Measles)                     -0.0689
 7 cor(Life.expectancy, Population)                  -0.0223
 8 cor(Life.expectancy, Total.expenditure)           0.175
 9 cor(Life.expectancy, Hepatitis.B)                 0.200
10 cor(Life.expectancy, Polio)                       0.327
11 cor(Life.expectancy, Diphtheria)                  0.341
12 cor(Life.expectancy, Alcohol)                     0.403
13 cor(Life.expectancy, percentage.expenditure)      0.410
14 cor(Life.expectancy, GDP)                         0.441
15 cor(Life.expectancy, BMI)                         0.542
16 cor(Life.expectancy, Income.composition.of.resources) 0.721
17 cor(Life.expectancy, Schooling)                   0.728
```

- Calculated correlation coefficient as our summary statistic

- Used this to choose Schooling, ICOR, and HIV/AIDs as top variables

- Plotted visualizations
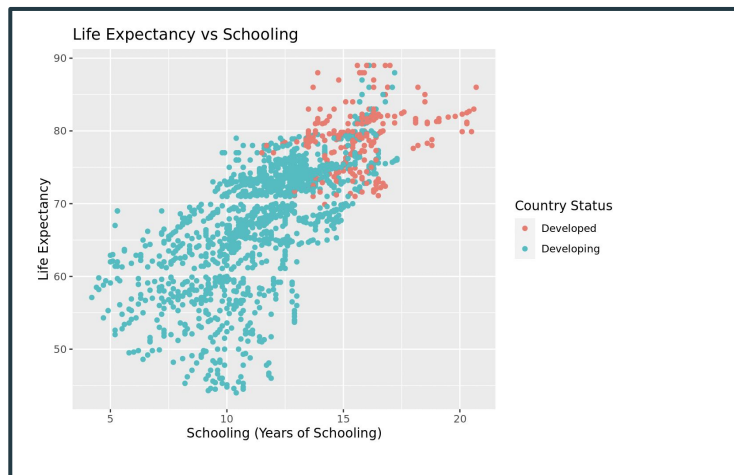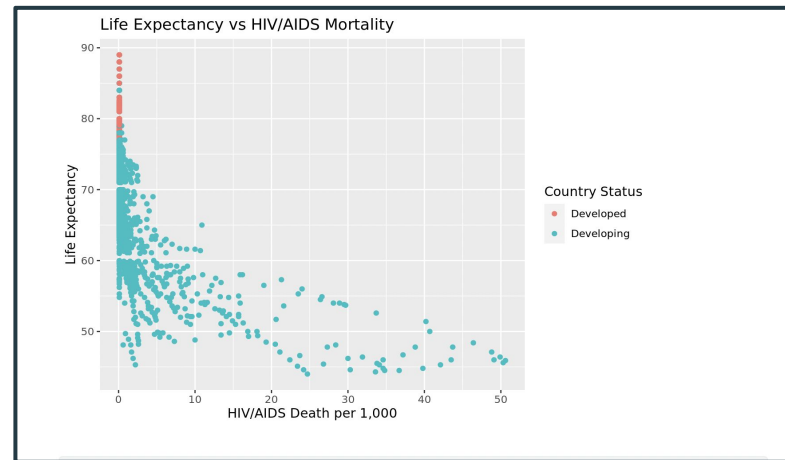
Additive model

- Based on the linear trends from the EDA, linear regression is appropriate
- AIC of 9464.471

$$Life{\widehat{Expectan}}cy = 48.56 + 1.195 * schooling - 0.63 * HIV/AIDS$$
$$+ 13.867 * ICOR + -1.45 * status$$

$$status = \begin{cases} 1 & \text{if Developing Country} \\ 0 & \text{if Developed Country} \end{cases}$$

Interactive model

- Added an interactive term on ICOR and development status
- Lower AIC of 9453.634

$$Life\widehat{Expect}ancy = 33.120 + 1.162 * schooling - 0.63 * HIV/AIDS$$
$$+ 32.927 * ICOR + 14.33 * status - 19.025 * ICOR * status$$

$$status = \begin{cases} 1 & \text{if Developing Country} \\ 0 & \text{if Developed Country} \end{cases}$$

Conclusions + Future Directions

- Limited scope of inference means we cannot conclude causation

- Provide avenues for future research

  - ex) more investigation needed into economic structures and the factors underlying the ICOR score

- More robust statistical analysis can inform policy making and governments