

Introduction to Regression Group Work

Alix Ziff, Gaea Morales, and Zachary Johnson

Summer 2020, Version 1: June 24

Introduction to Regression Analysis in R

In this group work, we are going to continue working with the Human Development Index data. This time we are asking: *Does a country's human development rank predict the infant mortality rate of its population?*

The data provided has each country's most recent ranking according to the Human Development Index. The goal will be to clean the data, make a scatterplot of the relationship, draw a line of best fit, and interpret some regression coefficients.

Clean our HDI Data

Exercise 1: Reshape the data into country-year format.

Visualize the Relationship

Exercise 2: Create a scatterplot visualizing the relationship between life expectancy and human development index ranking. Label the axes.

Exercise 3: If you have access to a printer, print your scatterplot. If not, copy a simplified (fewer observations) version of the scatterplot data with pen and paper.

Exercise 4: Draw an estimated line of best fit.

Interpret the Regression Results

Exercise 5: Label the predicted value, the actual value, and the error for three observations.

Bonus Exercise: Estimate the slope and the y intercept of their regression line, and explain what those mean.